RESULT 5
AF045938
LOCUS AF045938 777 bp DNA linear BCT 02-MAY-1998
DEFINITION Mycobacterium smegmatis putative ABC transporter nucleotide binding subunit (mtp1) gene, complete cds.
ACCESSION AF045938
VERSION AF045938.1 GI:3098527
KEYWORDS .
SOURCE Mycobacterium smegmatis.
ORGANISM Mycobacterium smegmatis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
REFERENCE 1 (bases 1 to 777)
AUTHORS Banerjee,S.K., Misra,P., Bhatt,K., Mande,S.C. and Chakraborti,P.K.
TITLE Identification of an ABC transporter gene that exhibits mRNA level overexpression in Fluoroquinolone resistant Mycobacterium smegmatis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Banerjee,S.K., Misra,P., Bhatt,K., Mande,S.C. and Chakraborti,P.K.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-1998) Institute of Microbial Technology, Sector 39-A, Chandigarh 160 036, India
FEATURES Location/Qualifiers
source 1. .777
/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:1772"
gene 1. .777
/gene="mtp1"
CDS 1. .777
/gene="mtp1"
/function="involved in fluoroquinolone resistance"
/note="PstB"
/codon_start=1
/transl_table=11
/product="putative ABC transporter nucleotide binding subunit"
/protein_id="AAC15686.1"
/db_xref="GI:3098528"
/translation="MAKRLDLKEVNIYYGGSTRLPTCHSRCSRAASRRIISQSGCGKS
TVLRTLNRMHEVIPGARVEGSVLLDGEDIYGPGVDPVGVRKTIGMVFQRPNPFPTMSI
RDNAVAGLKLQGVRNKKTLDEVAERSLRGANLWNEVKDRLDKPGGGLSGGQQQRLCIA
RAIAVQPDVLLMDDPCSALDPISTLAIEDLIATLKLDYTIVIVTHNMQQAARVSDQTA
FFNLEATGKPGRLIEIDDTEKIFSNPRQKATEDYISGRFG"
BASE COUNT 147 a 257 c 252 g 121 t
ORIGIN

Query Match 48.0%; Score 429.4; DB 1; Length 777;
Best Local Similarity 74.2%; Pred. No. 5.5e-89;
Matches 569; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

```
Qy    109 GCTCAAGCTCAATGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAA 168
          |||  | ||||| || ||||||||||||||||||| | |||||||| ||     ||||||
Db     12 GCTTGATCTCAAAGAGGTCAACATCTACTACGGCG-GTTCCACGCGGTTGCCGACGTGTC 70

Qy    169 CCTCGAGGTTCCTGCACGCTCTGTCACCGCAT-TCATCGGACCATCCGGCTGTGGCAAGT 227
```

Fri Apr 25 10:25:58 2003

u

```
Db     71 ACTCGCGGTGCAGCCGCGCAGCGTCACGGCGTATCATCAGCCAGTCGGGCTGCGGCAAGT 130
Qy    228 CCACAGTTCTCCGCTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGG 287
Db    131 CCACGGTGCTGCGCACGCTCAACCGCATGCACGAGGTCATCCCCGGCGCGCGCGTCGAGG 190
Qy    288 GCGAGATCCTTCTCGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCC 347
Db    191 GGTCGGTGCTGCTCGACGGCGAGGACATCTACGGGCCGGGAGTGGACCCGGTGGGTGTCC 250
Qy    348 GTAACACCATCGGCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGG 407
Db    251 GCAAGACCATCGGCATGGTGTTCCAGCGCCCGAACCCCTTCCCCACCATGTCGATTCGCG 310
Qy    408 ACAACGTGGTTGCAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAG 467
Db    311 ACAACGCCGTGGCCGGCCTCAAACTGCAGGGGGTGCGCAACAAGAAGACCCTCGACGAGG 370
Qy    468 TTGCTGAGAAGTCTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACA 527
Db    371 TCGCCGAGCGCTCACTGCGTGGTGCCAACCTGTGGAACGAGGTCAAGGACCGGCTCGACA 430
Qy    528 AGCCAGGCGGCGGCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCG 587
Db    431 AACCGGGCGGCGGGCTGTCGGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGTGCGATCG 490
Qy    588 CGGTTGAGCCAGAAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAA 647
Db    491 CGGTGCAGCCCGATGTGCTGCTGATGGACGACCCGTGCTCGGCGCTCGACCCGATCTCGA 550
Qy    648 CCCTGGCTGTGGAGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGA 707
Db    551 CGCTGGCGATCGAGGACCTGATCGCGACGCTCAAGCTCGACTACACGATCGTCATCGTCA 610
Qy    708 CCCACAACATGCAGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGG 767
Db    611 CGCACAACATGCAGCAGGCCGCGCGCGTGAGCGATCAGACCGCGTTCTTCAACCTCGAGG 670
Qy    768 CGACCGGTAGGCCAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACC 827
Db    671 CCACCGGAAAGCCGGGCCGGCTGATCGAGATCGACGACACCGAGAAGATCTTCTCCAATC 730
Qy    828 CAGATCAGAAGGAAACCGAAGATTACATCTCCGGCCGCTTCGGATAA 874
Db    731 CCCGCCAGAAGGCGACCGAGGATTACATCTCGGGCCGCTTCGGCTGA 777
```

RESULT 6

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2003, 14:54:56 ; Search time 2497 Seconds
(without alignments)
10419.667 Million cell updates/sec

Title: US-09-603-024C-1
Perfect score: 894
Sequence: 1 tcatcatcgctgtcctgaac.....atcgaagaattaaagccact 894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	892.4	99.8	320550	1	AP005282	AP005282 Corynebac
c 2	892.4	99.8	349980	6	AX127151	AX127151 Sequence
c 3	892.4	99.8	349980	6	AX127152	AX127152 Sequence
4	869.4	97.2	897	6	AX122914	AX122914 Sequence
5	429.4	48.0	777	1	AF045938	AF045938 Mycobacte
6	417.6	46.7	14029	1	AE006973	AE006973 Mycobacte
7	417.6	46.7	68848	1	MTV043	AL022004 Mycobacte
c 8	416.4	46.6	32354	1	SCD84	AL353816 Streptomy
9	412	46.1	9302	1	SGR243674	AJ243674 Streptomy
10	376.8	42.1	40123	1	MLU15182	U15182 Mycobacteri
c 11	376.8	42.1	342300	1	MLEPRTN8	AL583924 Mycobacte
12	375.2	42.0	37769	1	MSGL611CS	L78822 Mycobacteri
13	336.2	37.6	213732	1	AE001862	AE001862 Deinococc
14	289.2	32.3	3234	1	PSEPA1G	D28587 Pseudomonas
15	289.2	32.3	6125	1	PSEPSTD	D45195 Pseudomonas
c 16	289.2	32.3	11334	1	AE004948	AE004948 Pseudomon
17	282.8	31.6	7736	1	AB017356	AB017356 Pseudomon
c 18	282.6	31.6	349498	1	AP003002	AP003002 Mesorhizo
19	274.4	30.7	11092	1	AE005703	AE005703 Caulobact
20	267.4	29.9	202050	1	AL646065	AL646065 Ralstonia
21	264.6	29.6	10083	1	AE009013	AE009013 Agrobacte
22	264.6	29.6	11188	1	AE007979	AE007979 Agrobacte
23	258.6	28.9	8584	1	AB074879	AB074879 Rhodobact
24	257.8	28.8	2164	1	AF196490	AF196490 Caulobact
c 25	256.2	28.7	14375	1	AE009631	AE009631 Brucella
26	255	28.5	300000	1	SME591784	AL591784 Sinorhizo
27	253.4	28.3	12445	1	AE012855	AE012855 Chlorobiu
c 28	243	27.2	13861	1	AE000720	AE000720 Aquifex a
c 29	242.4	27.1	16155	1	AE004999	AE004999 Halobacte
30	240	26.8	13912	1	AE000929	AE000929 Methanoba
31	238.4	26.7	5678	1	D89963	D89963 Enterobacte
c 32	238	26.6	105795	1	SYCCPNC	D64001 Synechocyst
c 33	233.8	26.2	21130	1	AE008879	AE008879 Salmonell
34	233.8	26.2	245050	1	AL627280	AL627280 Salmonell
c 35	229.2	25.6	12627	1	AE012253	AE012253 Xanthomon
36	229	25.6	11140	1	AE004350	AE004350 Vibrio ch
37	224.2	25.1	3860	1	ECPHOWTU	X02723 E. coli DNA
38	224.2	25.1	5032	1	ECOPHOS	K01992 E.coli phos
c 39	224.2	25.1	12989	1	AE000449	AE000449 Escherich
c 40	224.2	25.1	136254	1	ECOUW82	L10328 E. coli; th
41	223	24.9	5225	1	AF248495	AF248495 Edwardsie
42	222.8	24.9	774	6	AX188983	AX188983 Sequence
c 43	222.6	24.9	11809	1	AE005603	AE005603 Escherich
c 44	222.6	24.9	280900	1	AP002566	AP002566 Escherich
c 45	219.6	24.6	10221	1	AE011789	AE011789 Xanthomon

ALIGNMENTS

RESULT 1
AP005282/c
LOCUS AP005282 320550 bp DNA linear BCT 08-AUG-2002
DEFINITION Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 9/10.
ACCESSION AP005282 BA000036
VERSION AP005282.1 GI:21325287
KEYWORDS .
SOURCE Corynebacterium glutamicum ATCC 13032 (strain:ATCC 13032) DNA.
ORGANISM Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Nakagawa,S.
TITLE Complete genomic sequence of Corynebacterium glutamicum ATCC 13032

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 320550)
  AUTHORS   Nakagawa,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.
            Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,
            Tokyo 194-8533, Japan (E-mail:snakagawa@xanagen.com,
            Tel:81-44-829-3031, Fax:81-44-813-1651)
COMMENT     This sequence is conducted by collaboration of Kyowa Hakko Kogyo
            Co. Ltd. And Kitasato University.
FEATURES             Location/Qualifiers
     source          1. .320550
                     /organism="Corynebacterium glutamicum ATCC 13032"
                     /strain="ATCC 13032"
                     /db_xref="taxon:196627"
                     /note="ATCC 13032"
     gene            188. .814
                     /gene="Cgl2516"
     CDS             188. .814
                     /gene="Cgl2516"
                     /codon_start=1
                     /transl_table=11
                     /product="Hypothetical protein"
                     /protein_id="BAB99909.1"
                     /db_xref="GI:21325288"
                     /translation="MYRAFDFVPLSMLVLVTKTRRKDHRMETLAAQARTLLEKWGVAP
                     THASFVESIAKAIPILSILLTLIVTVNGISSGNPVQPPALEQVRTDVVNKINYERNLK
                     GLVSISPELELHTAAQTIAQRNADSDSEEKVPDPEGNLVVLQQNLPYANANADTIVDR
                     FLNSPDHVKLLLANDYEAIGVGVAYKGDHAWIVVEFTVAPADSVESTE"
     gene            811. .2151
                     /gene="Cgl2517"
     CDS             811. .2151
                     /gene="Cgl2517"
                     /codon_start=1
                     /transl_table=11
                     /product="Nicotinic acid phosphoribosyltransferase"
                     /protein_id="BAB99910.1"
                     /db_xref="GI:21325289"
                     /translation="MNTNPSEFSSNRSTALLTDKYELTMLQAALADGSAERPSTFEVF
                     SRRLPNERRYGVVAGTARVLKAIRDFVFTEEQLADLDFLDDRTLEYLRNYRFTGQVDG
                     YREGEIYFPQSPLLTVRGTFAECVILETVILSIMNADSAVASAAARMVTAADGRPIIE
                     MGSRRTHEYSAVTASRAAYLAGFSTTSNLEAAYRYGIPASGTSAHAWTLLHINDDGTP
                     NEAAAFKAQVESLGVDTTLLVDTYDITQGVATAIEVAGPDLGGVRIDSGDLGVLARKV
                     RKQLDDLNAHNTKIVVSSDLDEFAIAGLRGEPVDVFGVGTSVVTGSGAPTAGLVYKIV
                     EVAGHPVAKRSRNKESYGGGKKAVRTHRKSGTAIEEIVYPFNAEAPDTGKLDTLSLTI
                     PLMRDGEIVPGLPTLEDSRAYLAKQLVSLPWEGLALSRDEPVLHTRFVGFPPAA"
     gene            2438. .2746
                     /gene="Cgl2518"
     CDS             2438. .2746
                     /gene="Cgl2518"
                     /codon_start=1
                     /transl_table=11
                     /product="Hypothetical protein"
                     /protein_id="BAB99911.1"
                     /db_xref="GI:21325290"
                     /translation="MGLRINLFGRPPQKTQITWKLPAQNRPQAPLSKQSNQTVNRCFA
                     WLSLSTPIWLDLAQCLRLNHPQIPPQIRSMLPRKSCLLLPLKRWAVPVVLDRKQWLRL
                     "
     gene            2866. .4608
                     /gene="Cgl2519"
     CDS             2866. .4608
                     /gene="Cgl2519"
                     /codon_start=1
                     /transl_table=11
                     /product="Rad3-related DNA helicases"
                     /protein_id="BAB99912.1"
                     /db_xref="GI:21325291"
                     /translation="MSTATIALQRQLVNRDLPRLVDALEPLMERRPTFAIMKGRSNYL
                     CMNKVARQEELNQEDALIEQEDISWLGKHIVRLNEWANETETGDRDDLDPGVPDLAWK
                     QVSVTARECIGASRCPHGEDCFAEIARGKAKEADVVVTNHALLAIDALSDVSVLPEHD
                     VVVIDEAHELDGRITAVASAEITVNSLNLAARRASKLDSDKREERVQEIAGDLETLLQ
                     TMQPGRWNDMDEGSKGTLVALKDALWALRAQIAGAPEGEAANDPERFAERQNLSNHLM
                     EIHDANVRILEVFAEEDPSKQYDVVWHNHDDRRGDSLNVAPLSVAGLLHEKLFAENTV
                     VLASATLTIGGNFNAMAASWGLPKGSWDSMDAGTPFDPAKSGILYTARHLPDPGRDGL
                     PEETLDEIYELITAAGGRTLGLFSSKRAAEQATKAMRLRLPFDVLCQGDDNTAALVKK
                     FSDSENTCLFGTLTLWQGVDVPGRSLSLVLIDRIPFPRPDDPLLQARKEAADAEGRNG
                     FMEVAATHAALLMAQGAGRLLRHVGDRGVVAVLDHRLSTKRYGGFLRFSMPRFWETTN
                     PETVRAALKRLVTK"
     gene            complement(4618. .5511)
                     /gene="Cgl2520"
     CDS             complement(4618. .5511)
                     /gene="Cgl2520"
                     /codon_start=1
                     /transl_table=11
                     /product="Hypothetical membrane protein"
                     /protein_id="BAB99913.1"
                     /db_xref="GI:21325292"
                     /translation="MTSPYGSQYPGDDNNNWNSQFGNPSGEQNYGQPYGAPYGQPYGQ
                     PFDQGFNAYSSPIPPEVPQPSMQEAQWRSFDLGTVFGQAWKGFTATWQAWVLSALIYF
                     AVLLVLMFAWILPMVSVLAATSSGSDSAAIAATGGTSFFGFMLMIVLAFISFVYSLNC
                     YRNAARVVRGEQITIQSFFKMKGLGKALGIYILINIVIFIGMILLLIPGIIAAVVLIF
                     AVPVAFQLRDASIGDAFSASWKAVSKNVGQVILLELAIFALSFLGSAVIIGMLVTTPL
                     TFLLYAYAFQTASGGPIMQRQ"
     gene            complement(5587. .6312)
                     /gene="Cgl2521"
     CDS             complement(5587. .6312)
                     /gene="Cgl2521"
                     /note="PF01981:Domain of unknown function UPF0099"
                     /codon_start=1
                     /transl_table=11
                     /product="Uncharacterized ACR"
                     /protein_id="BAB99914.1"
                     /db_xref="GI:21325293"
                     /translation="MPKDSLLDAHLLLKARASDGPHSEDPSDPSTVQAMQIALHIPKQ
                     NPPRRTDVLEAAARSVVKLCLDERVSTDPDFRAALERWYGHLIRKVSRRARNAAWDRV
                     QDLPGVTVEDDAAKVRAFLPSAVVDVPADIKKLQISGTELPLDEPNPINDEYPVIYID
                     ESLKMTLGKAAAQVGHASMLLAAHQPFEWVEQWEAADFALHVREIPSEEFLRLIESPG
                     AVPVRDGGFTEVAPNTVTVVAIP"
     gene            complement(6305. .7324)
                     /gene="Cgl2522"
     CDS             complement(6305. .7324)
                     /gene="Cgl2522"
                     /EC_number="3.1.3.3"
                     /note="PF00702:haloacid dehalogenase-like hydrolase
                     TIGR00099:TIGR00099: conserved hypothetical protein
                     TIGR00338:serB: phosphoserine phosphatase SerB"
                     /codon_start=1
                     /transl_table=11
                     /product="Phosphoserine phosphatase"
                     /protein_id="BAB99915.1"
                     /db_xref="GI:21325294"
                     /translation="MHGQSVVVELQETVQSSRPRSSHVVVVLGDPVDALDISRIGQTL
                     ADYDANIDTIRGISDYPVTGLELKVTVPDVSPGGGEAMRKALAALTSELNVDIAIERS
                     GLLRRSKRLVCFDCDSTLITGEVIEMLAAHAGKEAEVAAVTERAMRGELDFEESLRER
                     VKALAGLDASVIDEVAAAIELTPGARTTIRTLNRMGYQTAVVSGGFIQVLEGLAEELE
                     LDYVRANTLEIVDGKLTGNVTGKIVDRAAKAEFLREFAADSGLKMYQTVAVGDGANDI
                     DMLSAAGLGVAFNAKPALKEIADTSVNHPFLDEVLHIMGISRDEIDLADQEDGTFHRV
                     PLTNA"
     gene            complement(7811. .9565)
                     /gene="Cgl2523"
     CDS             complement(7811. .9565)
                     /gene="Cgl2523"
                     /EC_number="1.9.3.1"
                     /note="PF00115:Cytochrome C and Quinol oxidase polypeptide
                     I"
                     /codon_start=1
                     /transl_table=11
                     /product="Heme/copper-type cytochrome/quinol oxidases,
                     subunit 1"
                     /protein_id="BAB99916.1"
                     /db_xref="GI:21325295"
                     /translation="MTAVAPRVDGHVAPQRPEPTGHARKGSKAWLMMTTTDHKQLGIM
                     YIIMSFSFFFLGGLMALLIRAELFTPGLQFLSNEQFNQLFTMHGTVMLLLYGTPIVWG
                     FANYVLPLQIGAPDVAFPRLNAFGFWITTVGGVAMLTGFLTPGGAADFGWTMYSPLSD
                     AIHSPGLGSDMWIVGVGATGIGSVASAINMLTTILCLRAPGMTMFRMPIFTWNIFVVS
                     VLALLIFPLLLAAALGVLYDRKLGGHLYDPANGGSLLWQHLFWFFGHPEVYVLALPFF
```

```
                     GIVSEIIPVFSRKPMFGYVGLIFATLSIGALSMAVWAHHMFVTGAVLLPFFSFMTFLI
                     SVPTGVKFFNWVGTMWKGHITWETPMIWSVGFMATFLFGGLTGIMLASPPLDFHLADS
                     YFLIAHFHYTLFGTVVFASCAGVYFWFPKMTGRMMDERLGKIHFWLTFVGFHGTFLIQ
                     HWVGNMGMPRRYADYLDSDGFTIYNQISTVFSFLLGLSVIPFIWNVFKSWRYGELVTV
                     DDPWGYGNSLEWATSCPPPRHNFASLPRIRSERPAFELHYPHMIERMRAEAHTGHHDD
                     INAPELGTAPALASDSSR"
     gene            9701. .10009
                     /gene="Cgl2524"
     CDS             9701. .10009
                     /gene="Cgl2524"
                     /codon_start=1
                     /transl_table=11
                     /product="Hypothetical protein"
                     /protein_id="BAB99917.1"
                     /db_xref="GI:21325296"
                     /translation="MPKKNVFQTIHEVAARFSHLWAVTFNQRCEGVHNLTDISPGSPT
                     PTVNPCHRFHNSPRRGSIEWLRRPFQGRFLLKRDVPPRGPRGAWRSPTPICASEKPAP
                     "
     gene            complement(10086. .11090)
                     /gene="Cgl2525"
     CDS             complement(10086. .11090)
                     /gene="Cgl2525"
                     /note="PF00268:Ribonucleotide reductase, small chain"
                     /codon_start=1
                     /transl_table=11
                     /product="Ribonucleotide reductase beta subunit"
                     /protein_id="BAB99918.1"
                     /db_xref="GI:21325297"
                     /translation="MAADSDLSVHDAYLKEHVAPVKAINWNSIPDSKDLEVWDRLTGN
                     FWLPEKVPVSNDIKSWGTLNEVEKAATMRVFTGLTLLDTIQGTVGAISLLPDADSLHE
                     EAVLTNIAFMESVHAKSYSNIFMTLASTAEINDAFRWSEENENLQRKAKIILSYYEGD
                     DPLKRKIASVILESFLFYSGFYLPMYWSSHSKLTNTADVIRLIIRDEAVHGYYIGYKY
                     QKAVAKETPERQEELKEYTFDLLYDLYDNETQYSEDLYDDLGWTEDVKRFLRYNANKA
                     LNNLGYEGLFPADETKVSPNILSALSPNADENHDFFSGSGSSYVIGKAENTEDDDWDF
                     "
     gene            11555. .12043
                     /gene="Cgl2526"
     CDS             11555. .12043
                     /gene="Cgl2526"
                     /note="PF00210:Ferritin
                     TIGR00754:bfr: bacterioferritin"

  Query Match             99.8%;  Score 892.4;  DB 1;  Length 320550;
  Best Local Similarity   99.9%;  Pred. No. 6.4e-196;
  Matches  893;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TCATCATCGCTGTCCTGAACATTGGCGCACGAATCATCTCCGCCAAGTTCTCTGTCAAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71076 TCATCATCGCTGTCCTGAACATTGGCGCACGAATCATCTCCGCCAAGTTCTCTGTCAAGC 71017

Qy      61 AATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGATGTCGAAGCTCAAGCTCAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71016 AATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGATGTCGAAGCTCAAGCTCAA 70957

Qy     121 TGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70956 TGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCC 70897

Qy     181 TGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70896 TGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCG 70837

Qy     241 CTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70836 CTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCT 70777

Qy     301 CGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGG 360
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70776 CGACGGTGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGG 70717

Qy     361 CATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70716 CATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGC 70657

Qy     421 AGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70656 AGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTC 70597

Qy     481 TCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70596 TCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGG 70537

Qy     541 CCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70536 CCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGA 70477

Qy     601 AATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70476 AATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGA 70417

Qy     661 GGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70416 GGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCA 70357

Qy     721 GCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70356 GCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCC 70297

Qy     781 AGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70296 AGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGA 70237

Qy     841 AACCGAAGATTACATCTCCGGCCGCTTCGGATAAATCGAAGAATTAAAGCCACT 894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70236 AACCGAAGATTACATCTCCGGCCGCTTCGGATAAATCGAAGAATTAAAGCCACT 70183


RESULT 2
AX127151/c
LOCUS       AX127151               349980 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 7067 from Patent EP1108790.
ACCESSION   AX127151 AX114121
VERSION     AX127151.1  GI:14041139
KEYWORDS    .
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 349980)
  AUTHORS   Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
  TITLE     Novel polynucleotides
  JOURNAL   Patent: EP 1108790-A 7067 20-JUN-2001;
            KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
                     /note="Seq 1 to long (3.309.400) split in 11, seq 7067
                     2.400.001 2.749.980"
BASE COUNT    80900 a  98397 c  92139 g  78544 t
ORIGIN

  Query Match             99.8%;  Score 892.4;  DB 6;  Length 349980;
  Best Local Similarity   99.9%;  Pred. No. 6.5e-196;
  Matches  893;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TCATCATCGCTGTCCTGAACATTGGCGCACGAATCATCTCCGCCAAGTTCTCTGTCAAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 334325 TCATCATCGCTGTCCTGAACATTGGCGCACGAATCATCTCCGCCAAGTTCTCTGTCAAGC 334266

Qy      61 AATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGATGTCGAAGCTCAAGCTCAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 334265 AATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGATGTCGAAGCTCAAGCTCAA 334206
```

```
Qy    121 TGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 334205 TGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCC 334146

Qy    181 TGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 334145 TGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCG 334086

Qy    241 CTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 334085 CTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCT 334026

Qy    301 CGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGG 360
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 334025 CGACGGTGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGG 333966

Qy    361 CATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333965 CATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGC 333906

Qy    421 AGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333905 AGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTC 333846

Qy    481 TCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333845 TCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGG 333786

Qy    541 CCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333785 CCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGA 333726

Qy    601 AATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333725 AATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGA 333666

Qy    661 GGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333665 GGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCA 333606

Qy    721 GCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333605 GCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCC 333546

Qy    781 AGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333545 AGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGA 333486

Qy    841 AACCGAAGATTACATCTCCGGCCGCTTCGGATAAATCGAAGAATTAAAGCCACT 894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 333485 AACCGAAGATTACATCTCCGGCCGCTTCGGATAAATCGAAGAATTAAAGCCACT 333432
```

RESULT 3
AX127152/c

```
LOCUS       AX127152              349980 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 7068 from Patent EP1108790.
ACCESSION   AX127152 AX114121
VERSION     AX127152.1  GI:14041140
KEYWORDS    .
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 349980)
  AUTHORS   Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
  TITLE     Novel polynucleotides
  JOURNAL   Patent: EP 1108790-A 7068 20-JUN-2001;
            KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
                     /note="Seq 1 to long (3.309.400) split in 11, seq 7068
                     2.700.001 3.049.980"
BASE COUNT    81250 a  97718 c  90621 g  80391 t
ORIGIN
```

Query Match 99.8%; Score 892.4; DB 6; Length 349980;
Best Local Similarity 99.9%; Pred. No. 6.5e-196;
Matches 893; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 TCATCATCGCTGTCCTGAACATTGGCGCACGAATCATCTCCGCCAAGTTCTCTGTCAAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34325 TCATCATCGCTGTCCTGAACATTGGCGCACGAATCATCTCCGCCAAGTTCTCTGTCAAGC 34266

Qy     61 AATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGATGTCGAAGCTCAAGCTCAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34265 AATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGATGTCGAAGCTCAAGCTCAA 34206

Qy    121 TGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34205 TGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCC 34146

Qy    181 TGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34145 TGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCG 34086

Qy    241 CTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34085 CTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCT 34026

Qy    301 CGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGG 360
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34025 CGACGGTGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGG 33966

Qy    361 CATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33965 CATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGC 33906

Qy    421 AGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33905 AGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTC 33846

Qy    481 TCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33845 TCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGG 33786

Qy    541 CCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33785 CCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGA 33726

Qy    601 AATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33725 AATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGA 33666

Qy    661 GGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33665 GGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCA 33606

Qy    721 GCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33605 GCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCC 33546

Qy    781 AGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33545 AGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGA 33486

Qy    841 AACCGAAGATTACATCTCCGGCCGCTTCGGATAAATCGAAGAATTAAAGCCACT 894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  33485 AACCGAAGATTACATCTCCGGCCGCTTCGGATAAATCGAAGAATTAAAGCCACT 33432
```

RESULT 4

```
AX122914
LOCUS       AX122914                897 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 2830 from Patent EP1108790.
ACCESSION   AX122914
VERSION     AX122914.1  GI:14040402
KEYWORDS    .
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
  TITLE     Novel polynucleotides
  JOURNAL   Patent: EP 1108790-A 2830 20-JUN-2001;
            KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES             Location/Qualifiers
     source          1. .897
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
BASE COUNT      221 a    262 c    231 g    183 t
ORIGIN

  Query Match             97.2%;  Score 869.4;  DB 6;  Length 897;
  Best Local Similarity   99.9%;  Pred. No. 6.9e-191;
  Matches  870;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TCATCATCGCTGTCCTGAACATTGGCGCACGAATCATCTCCGCCAAGTTCTCTGTCAAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 TCATCATCGCTGTCCTGAACATTGGCGCACGAATCATCTCCGCCAAGTTCTCTGTCAAGC 86

Qy      61 AATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGATGTCGAAGCTCAAGCTCAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87 AATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGATGTCGAAGCTCAAGCTCAA 146

Qy     121 TGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147 TGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCC 206

Qy     181 TGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     207 TGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCG 266

Qy     241 CTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     267 CTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCT 326

Qy     301 CGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGG 360
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     327 CGACGGTGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGG 386

Qy     361 CATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     387 CATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGC 446

Qy     421 AGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     447 AGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTC 506

Qy     481 TCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     507 TCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGG 566

Qy     541 CCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     567 CCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGA 626

Qy     601 AATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     627 AATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGA 686

Qy     661 GGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCA 746

Qy     721 GCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     747 GCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCC 806

Qy     781 AGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     807 AGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGA 866

Qy     841 AACCGAAGATTACATCTCCGGCCGCTTCGGA 871
          |||||||||||||||||||||||||||||||
Db     867 AACCGAAGATTACATCTCCGGCCGCTTCGGA 897


RESULT 5
AF045938
LOCUS       AF045938                777 bp    DNA     linear   BCT 02-MAY-1998
DEFINITION  Mycobacterium smegmatis putative ABC transporter nucleotide binding
            subunit (mtp1) gene, complete cds.
ACCESSION   AF045938
VERSION     AF045938.1  GI:3098527
KEYWORDS    .
SOURCE      Mycobacterium smegmatis.
  ORGANISM  Mycobacterium smegmatis
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium.
REFERENCE   1  (bases 1 to 777)
  AUTHORS   Banerjee,S.K., Misra,P., Bhatt,K., Mande,S.C. and Chakraborti,P.K.
  TITLE     Identification of an ABC transporter gene that exhibits mRNA level
            overexpression in Fluoroquinolone resistant Mycobacterium smegmatis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 777)
  AUTHORS   Banerjee,S.K., Misra,P., Bhatt,K., Mande,S.C. and Chakraborti,P.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-1998) Institute of Microbial Technology, Sector
            39-A, Chandigarh 160 036, India
FEATURES             Location/Qualifiers
     source          1. .777
                     /organism="Mycobacterium smegmatis"
                     /strain="mc2155"
                     /db_xref="taxon:1772"
     gene            1. .777
                     /gene="mtp1"
     CDS             1. .777
                     /gene="mtp1"
                     /function="involved in fluoroquinolone resistance"
                     /note="PstB"
                     /codon_start=1
                     /transl_table=11
                     /product="putative ABC transporter nucleotide binding
                     subunit"
                     /protein_id="AAC15686.1"
                     /db_xref="GI:3098528"
                     /translation="MAKRLDLKEVNIYYGGSTRLPTCHSRCSRAASRRIISQSGCGKS
                     TVLRTLNRMHEVIPGARVEGSVLLDGEDIYGPGVDPVGVRKTIGMVFQRPNPFPTMSI
                     RDNAVAGLKLQGVRNKKTLDEVAERSLRGANLWNEVKDRLDKPGGGLSGGQQQRLCIA
                     RAIAVQPDVLLMDDPCSALDPISTLAIEDLIATLKLDYTIVIVTHNMQQAARVSDQTA
                     FFNLEATGKPGRLIEIDDTEKIFSNPRQKATEDYISGRFG"
BASE COUNT      147 a    257 c    252 g    121 t
ORIGIN

  Query Match             48.0%;  Score 429.4;  DB 1;  Length 777;
  Best Local Similarity   74.2%;  Pred. No. 5.5e-89;
  Matches  569;  Conservative    0;  Mismatches  196;  Indels    2;  Gaps    2;

Qy     109 GCTCAAGCTCAATGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAA 168
           ||| | ||||| || ||||||||||||||||| | |||||||| ||     |||||
Db      12 GCTTGATCTCAAAGAGGTCAACATCTACTACGGCG-GTTCCACGCGGTTGCCGACGTGTC 70

Qy     169 CCTCGAGGTTCCTGCACGCTCTGTCACCGCAT-TCATCGGACCATCCGGCTGTGGCAAGT 227
```

```
Db     71 ACTCGCGGTGCAGCCGCGCAGCGTCACGGCGTATCATCAGCCAGTCGGGCTGCGGCAAGT 130

Qy    228 CCACAGTTCTCCGCTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGG 287

Db    131 CCACGGTGCTGCGCACGCTCAACCGCATGCACGAGGTCATCCCCGGCGCGCGCGTCGAGG 190

Qy    288 GCGAGATCCTTCTCGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCC 347

Db    191 GGTCGGTGCTGCTCGACGGCGAGGACATCTACGGGCCGGGAGTGGACCCGGTGGGTGTCC 250

Qy    348 GTAACACCATCGGCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGG 407

Db    251 GCAAGACCATCGGCATGGTGTTCCAGCGCCCGAACCCCTTCCCCACCATGTCGATTCGCG 310

Qy    408 ACAACGTGGTTGCAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAG 467

Db    311 ACAACGCCGTGGCCGGCCTCAAACTGCAGGGGGTGCGCAACAAGAAGACCCTCGACGAGG 370

Qy    468 TTGCTGAGAAGTCTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACA 527

Db    371 TCGCCGAGCGCTCACTGCGTGGTGCCAACCTGTGGAACGAGGTCAAGGACCGGCTCGACA 430

Qy    528 AGCCAGGCGGCGGCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCG 587

Db    431 AACCGGGCGGCGGGCTGTCGGGTGGTCAGCAGCAGCGTCTGTGCATCGCGCGTGCGATCG 490

Qy    588 CGGTTGAGCCAGAAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAA 647

Db    491 CGGTGCAGCCCGATGTGCTGCTGATGGACGACCCGTGCTCGGCGCTCGACCCGATCTCGA 550

Qy    648 CCCTGGCTGTGGAGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGA 707

Db    551 CGCTGGCGATCGAGGACCTGATCGCGACGCTCAAGCTCGACTACACGATCGTCATCGTCA 610

Qy    708 CCCACAACATGCAGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGG 767

Db    611 CGCACAACATGCAGCAGGCCGCGCGCGTGAGCGATCAGACCGCGTTCTTCAACCTCGAGG 670

Qy    768 CGACCGGTAGGCCAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACC 827

Db    671 CCACCGGAAAGCCGGGCCGGCTGATCGAGATCGACGACACCGAGAAGATCTTCTCCAATC 730

Qy    828 CAGATCAGAAGGAAACCGAAGATTACATCTCCGGCCGCTTCGGATAA 874

Db    731 CCCGCCAGAAGGCGACCGAGGATTACATCTCGGGCCGCTTCGGCTGA 777


RESULT 6
AE006973
LOCUS       AE006973               14029 bp    DNA     linear   BCT 27-APR-2001
DEFINITION  Mycobacterium tuberculosis CDC1551,  section 59 of 280 of the
            complete genome.
ACCESSION   AE006973 AE000516
VERSION     AE006973.1  GI:13880378
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis CDC1551.
  ORGANISM  Mycobacterium tuberculosis CDC1551
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE   1  (bases 1 to 14029)
  AUTHORS   Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
  TITLE     Whole genome comparison of Mycobacterium tuberculosis clinical and
            laboratory strains
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 14029)
  AUTHORS   Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2001) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .14029
                     /organism="Mycobacterium tuberculosis CDC1551"
                     /strain="CDC1551"
                     /db_xref="taxon:83331"
                     /note="clinical strain"
     gene            complement(74. .256)
                     /gene="MT0831"
     CDS             complement(74. .256)
                     /gene="MT0831"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK45073.1"
                     /db_xref="GI:13880379"
                     /translation="MGRGRAKAKQTKVARELKYSSPQTDFQRLQRELSGTGTDRLDGD
                     GPSDDDSWNDEDDWRR"
     gene            complement(403. .1509)
                     /gene="MT0832"
     CDS             complement(403. .1509)
                     /gene="MT0832"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAK45074.1"
                     /db_xref="GI:13880380"
                     /translation="MAAVPAPDPGPDAGAIWHYGDPLGEQRAGQADAVLVDRSHRAVL
                     TLDGGDRQTWLHSISTQHVSDLPEGASTQNLSLDGQGRVEDHWIQTELGGTTYLDTEP
                     WRGEPLLAYLRKMVFWSMVTPRAADMAVLSLLGPRLAEERVLDALGLDVLPAEWLAVP
                     LAGGGIVRRMPDGLAGQIELDVVVKRGDRADWQRRLTQAGVRPAGIWAYEAHRVAHRV
                     PARRPRLGVDTDERTIPHEVGWIGGPGAGAVHLNKGCYRGQETVARVHNLGRPPRMLV
                     LLHLDESVQRPSTGDAVLAGGRTVGRLGTVVEHVELGPVALALLKRGLPGDTALVTGP
                     EAEVAAVIDVDSLPPADDVGAGRRAVERLRGGIR"
     gene            1574. .2461
                     /gene="MT0833"
     CDS             1574. .2461
                     /gene="MT0833"
                     /note="similar to PID:857561 SP:P54692; identified by
                     sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="aminotransferase, class IV"
                     /protein_id="AAK45075.1"
                     /db_xref="GI:13880381"
                     /translation="MLRQTGVVVTLDGEILQPGMPLLHADDLAAVRGDGVFETLLVRD
                     GRACLVEAHLQRLTQSARLMDLPEPDLPRWRRAVEVATQRWVASTADEGALRLIYSRG
                     REGGSAPTAYVMVSPVPARVIGARRDGVSAITLDRGLPADGGDAMPWLMASAKTLSYA
                     VNMAVLRHAARQGAGDVIFVSTDGYVLEGPRSTVVIATDGDQGGGNPCLLTPPPWYPI
                     LRGTTQQALFEVARAKGYDCDYRALRVADLFDSQGIWLVSSMTLAARVHTLDGRRLPR
                     TPIAEVFAELVDAAIVSDR"
     gene            complement(2507. .3187)
                     /gene="MT0834"
     CDS             complement(2507. .3187)
                     /gene="MT0834"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAK45076.1"
                     /db_xref="GI:13880382"
                     /translation="MSSGAGSDATGAGGVHAAGSGDRAVAAAVERAKATAARNIPAFD
                     DLPVPADTANLREGADLNNALLALLPLVGVWRGEGEGRGPDGDYRFGQQIVVSHDGGD
                     YLNWESRSWRLTATGDYQEPGLREAGFWRFVADPYDPSESQAIELLLAHSAGYVELFY
                     GRPRTQSSWELVTDALARSRSGVLVGGAKRLYGIVEGGDLAYVEERVDADGGLVPHLS
                     ARLSRFVG"
```

```
gene            complement(3184. .3282)
                /gene="MT0835"
CDS             complement(3184. .3282)
                /gene="MT0835"
                /note="identified by Glimmer2; putative"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="AAK45077.1"
                /db_xref="GI:13880383"
                /translation="MVLFFEIMLVLATVVISWFALYTLYRLVTDES"
gene            complement(3350. .3652)
                /gene="MT0836"
CDS             complement(3350. .3652)
                /gene="MT0836"
                /note="identified by Glimmer2; putative"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAK45078.1"
                /db_xref="GI:13880384"
                /translation="MCSGPKQGLTLPASVDLEKETVITGRVVDGDGQAVGGAFVRLLD
                SSDEFTAEVVASATGDFRFFAAPGSWTLRALSAAGNGDAVVQPSGAGIHEVDVKIT"
gene            complement(3654. .4487)
                /gene="MT0837"
CDS             complement(3654. .4487)
                /gene="MT0837"
                /note="similar to GB:M29612 SP:P16385 PID:152687;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="thiosulfate sulfurtransferase"
                /protein_id="AAK45079.1"
                /db_xref="GI:13880385"
                /translation="MARCDVLVSADWAESNLHAPKVVFVEVDEDTSAYDRDHIAGAIK
                LDWRTDLQDPVKRDFVDAQQFSKLLSERGIANEDTVILYGGNNNWFAAYAYWYFKLYG
                HEKVKLLDGGRKKWELDGRPLSSDPVSRPVTSYTASPPDNTIRAFRDEVLAAINVKNL
                IDVRSPDEFSGKILAPAHLPQEQSQRPGHIPGAINVPWSRAANEDGTFKSDEELAKLY
                ADAGLDNSKETIAYCRIGERSSHTWFVLRELLGHQNVKNYDGSWTEYGSLVGAPIELG
                S"
gene            complement(4780. .5202)
                /gene="MT0838"
CDS             complement(4780. .5202)
                /gene="MT0838"
                /note="similar to PID:699219; identified by sequence
                similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="thioredoxin ThiX, putative"
                /protein_id="AAK45080.1"
                /db_xref="GI:13880386"
                /translation="MTTMIVASVATGALATIARWLLTRRSVILREVGPETTPAAPART
                AELGLSGAGPTVVHFRAPGCAPCDRVRRGVGDVCADLGDVAHIEVDLDSNPQAARRFS
                VLSLPTTLIFDVDGRQRYRTSGVPKAADLRSALKPLLA"
gene            complement(5199. .6011)
                /gene="MT0839"
CDS             complement(5199. .6011)
                /gene="MT0839"
                /note="identified by Glimmer2; putative"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAK45081.1"
                /db_xref="GI:13880387"
                /translation="MPMRKVLVGVTGAAIVVAVLIVGAVGADFGASIYAEYRLSTTVR
                KAANLRSDPFVAILRFPFIPQAMREHYAELEIKAFAVEHAGSGTATLEATMHSIDLSY
                ASWLIRPDAKLPVGELESRIIIDSMHLGRYLGISDLMVAAPRQESNDATGGTTESGIS
                GSRGLVFSGTPISANFAHRVSVLVDLSVASDDRATLVITPTAVVTGPDTADQPVPDDK
                RDAVLHAFASKLPNQKLPFGVVPNTVGARGSDVIIEGITRGVTISLDEFKQS"
gene            6102. .6908
                /gene="MT0840"
CDS             6102. .6908
                /gene="MT0840"
                /note="similar to GB:L03213 SP:Q05943 PID:153283;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="transcriptional regulator"
                /protein_id="AAK45082.1"
                /db_xref="GI:13880388"
                /translation="MNWETPFPRRWRGLLELLLLTSELYPDPVLPALSLLPHTVRTAP
                AEASSLLEAGNADAVLVDARNDLSSGRGLCRLLSSTGRSIPVLAVVSEGGLVAVSADW
                GLDEILLLSTGPAEIDARLRLVVGRRGDLADQESLGKVSLGELVIDEGTYTARLRGRP
                LDLTYKEFELLKYLAQHAGRVFTRAQLLHEVWGYDFFGGTRTVDVHVRRLRAKLGPEH
                EALIGTVRNVGYKAVRPARGRPPAADPDDEDADPGRDGMQEPLVDPLRSQ"
gene            6905. .7852
                /gene="MT0841"
CDS             6905. .7852
                /gene="MT0841"
                /note="identified by Glimmer2; putative"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAK45083.1"
                /db_xref="GI:13880389"
                /translation="MTALDWRSALTADEQRSVRALVTATTAVDGVAPVGEQVLRELGQ
                QRTEHLLVAGSRPGGPIIGYLNLSPPRGAGGAMAELVVHPQSRRRGIGTAMARAALAK
                TAGRNQFWAHGTLDPARATASALGLVGVRELIQMRRPLRDIPEPTIPDGVVIRTYAGT
                SDDAELLRVNNAAFAGHPEQGGWTAVQLAERRGEAWFDPDGLILAFGDSPRERPGRLL
                GFHWTKVHPDHPGLGEVYVLGVDPAAQRRGLGQMLTSIGIVSLARRLGGRKTLDPAVE
                PAVLLYVESDNVAAVRTYQSLGFTTYSVDTAYALAGTDN"
gene            7895. .8671
                /gene="MT0842"
CDS             7895. .8671
```

Query Match 46.7%; Score 417.6; DB 1; Length 14029;
Best Local Similarity 68.9%; Pred. No. 4.1e-86;
Matches 573; Conservative 0; Mismatches 259; Indels 0; Gaps 0;

```
Qy     42 GCCAAGTTCTCTGTCAAGCAATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGA 101
          ||   | |  ||| |     ||  |   | |  | ||  |   || || | || ||| |
Db   7839 GCACGGATAACTGACCGAAGATGTTCCCCCCCAAGAAGTCGTAAGCAGGAGCTTAAGTGG 7898

Qy    102 TGTCGAAGCTCAAGCTCAATGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGA 161
            | | |  |||| || ||||||||||||||||||    || || || ||
Db   7899 CCAAGCGGTTGGACCTCACGGACGTCAACATCTACTACGGGTCATTTCATGCGGTCGCTG 7958

Qy    162 ACGTGAACCTCGAGGTTCCTGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTG 221
          | |||   || | | |||   | |||   ||||| || |||||||| || || ||||| |
Db   7959 ATGTGTCGCTGGCGATTCTGCCCCGCAGCGTCACGGCGTTCATCGGTCCCTCGGGCTGCG 8018

Qy    222 GCAAGTCCACAGTTCTCCGCTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACG 281
          ||||| | || || || ||| || | || || ||||| ||||||| ||| || ||    |
Db   8019 GCAAGACGACGGTGCTGCGCACCTTGAACCGGATGCATGAGGTCATCCCCGGAGCTCGAG 8078

Qy    282 TCAAGGGCGAGATCCTTCTCGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTG 341
          || |||| |   | || ||||| |   |  | |||||||| | ||   |||||||| || |
Db   8079 TCGAGGGTGCCGTACTGCTCGATGATCAAGATATCTACGCCCCCGGTATCGACCCGGTCG 8138

Qy    342 CAGTCCGTAACACCATCGGCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCA 401
            |||||     | ||||| ||||| || |||  | | || ||||||||  |||||||| |
Db   8139 GTGTCCGCCGGGCAATCGGGATGGTGTTTCAGCGGCCGAATCCATTCCCCGCCATGTCGA 8198

Qy    402 TCGAGGACAACGTGGTTGCAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCA 461
          |    |||| |||||||| || ||||||||    || | |   ||    |||  ||||
Db   8199 TTCGCAACAATGTGGTTGCCGGCCTGAAGCTGCAGGGTGTGCGCAATCGCAAGGTGCTCG 8258

Qy    462 AGGAAGTTGCTGAGAAGTCTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTC 521
          | ||    || ||    || || || ||||||||||||||||||| ||||| |||||||| |
Db   8259 ACGATACGGCCGAATCCTCGCTGCGCGGCGCAAACCTGTGGGACGAGGTCAAGGATCGAC 8318

Qy    522 TGGACAAGCCAGGCGGCGGCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCG 581
          |||| || || |||||||||  | || || || ||||||||||  |||||||||| || |
Db   8319 TGGATAAACCCGGCGGCGGATTGTCTGGGGGGCAGCAGCAGCGGTTGTGCATCGCACGGG 8378

Qy    582 CGATCGCGGTTGAGCCAGAAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAA 641
```

```
                | ||||| ||  | || ||  |  | || ||||||||||| |||||| |||| |||||||
Db    8379 CAATCGCCGTGCAACCCGACGTGTTGCTGATGGACGAGCCCTGCTCCTCGCTGGACCCAA 8438

Qy     642 TCTCAACCCTGGCTGTGGAGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCA 701
           |||| ||| ||||  | || ||||| |||  |||||| ||| |  ||| ||||||||||
Db    8439 TCTCCACCATGGCCATCGAAGACCTGATCAGCGAGCTCAAGCAGCAGTACACCATCGTCA 8498

Qy     702 TCGTGACCCACAACATGCAGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCC 761
           |||| ||||| ||||||||||||||||| || |||   |||||||| || |||| |. ||
Db    8499 TCGTCACCCATAACATGCAGCAGGCTGCCCGGGTGAGTGATCAGACGGCATTCTTCAACC 8558

Qy     762 TGGAGGCGACCGGTAGGCCAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCG 821
           |||| |||   || | ||| || || ||||| || ||||     ||| |||| ||||||
Db    8559 TGGAAGCGGTGGGAAAGCCGGGGCGGCTGGTAGAGATCGCCAGCACCGAGAAAATCTTCT 8618

Qy     822 AAAACCCAGATCAGAAGGAAACCGAAGATTACATCTCCGGCCGCTTCGGATA 873
           |||||  | |||||||  ||||| || |||||||||| ||||||||  ||
Db    8619 CCAACCCGAACCAGAAGGCCACCGAGGACTACATCTCCGGGCGCTTCGGCTA 8670


RESULT 7
MTV043
LOCUS       MTV043                  68848 bp    DNA     linear   BCT 03-AUG-2001
DEFINITION  Mycobacterium tuberculosis H37Rv complete genome; segment 40/162.
ACCESSION   AL022004 AL123456
VERSION     AL022004.1  GI:3261550
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis H37Rv.
  ORGANISM  Mycobacterium tuberculosis H37Rv
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE   1  (bases 1 to 68848)
  AUTHORS   Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
            Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,
            Tekaia,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
            Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,
            Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
            Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
            Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,
            Squares,S., Sqares,R., Sulston,J.E., Taylor,K., Whitehead,S. and
            Barrell,B.G.
  TITLE     Deciphering the biology of Mycobacterium tuberculosis from the
            complete genome sequence
  JOURNAL   Nature 393 (6685), 537-544 (1998)
  MEDLINE   98295987
   PUBMED   9634230
REFERENCE   2  (bases 1 to 68848)
  AUTHORS   Parkhill,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
            tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
            Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
            Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
            75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
COMMENT     On Jun 27, 1998 this sequence version replaced gi:2916866.
            Notes:
            Details of M. tuberculosis sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have
            been renumbered from the original cosmid submissions but the old
            gene designations are in brackets after the new gene numbers.
            Gene prediction was based on a Hidden Markov Model of TB genes
            implemented in TBparse (Krogh) supplemented with visual inspection
            of positional base preference in codons, especially where there is
            an increase in the observed/expected third position G + C.
            CAUTION:  In some cases we may not have predicted the correct
            initiation codon.  Where possible we choose an initiation codon
            (atg, gtg, or ttg) which is preceded by an upstream ribosome
            binding site sequence (optimally 5-13bp before the initiation
            codon).  If this cannot be identified we choose the most upstream
            initiation codon.
FEATURES             Location/Qualifiers
     source          1. .68848
                     /organism="Mycobacterium tuberculosis H37Rv"
                     /strain="H37Rv"
                     /db_xref="taxon:83332"
     source          <1. .40
                     /organism="Mycobacterium tuberculosis H37Rv"
                     /strain="H37Rv"
                     /db_xref="taxon:83332"
                     /clone="Y7H7A"
     source          67806. .>68848
                     /organism="Mycobacterium tuberculosis H37Rv"
                     /strain="H37Rv"
                     /db_xref="taxon:83332"
                     /clone="Y31"
     misc_feature    <1. .67906
                     /note="fragment designated v043. Does not represent a
                     physical clone"
     gene            31. .1125
                     /gene="purM"
     CDS             31. .1125
                     /gene="purM"
                     /note="Rv0809, (MTV043.01), len: 364. purM, probable
                     5'-phosphoribosyl-5-aminoimidazole synthetase, similar to
                     many e.g. P12043|PUR5_BACSU (346 aa) and Salmonella
                     typhimurium U68765|STU68765_2 (345 aa). FASTA scores:
                     sp|P12043|PUR5_BACSU PHOSPHORIBOSYLFORMYLGLYCINAMIDINE
                     (346 aa)  opt: 1023 z-score: 1250.3 E(): 0; 46.5% identity
                     in 331 aa overlap; andU68765|STU68765_2  (345 aa)  opt:
                     1014 z-score: 1239.4 E():0;    47.6% identity in 330 aa
                     overlap. TBparse score is 0.901"
                     /codon_start=1
                     /transl_table=11
                     /product="purM"
                     /protein_id="CAA17615.1"
                     /db_xref="GI:2916867"
                     /db_xref="SPTREMBL:O53823"
                     /translation="MTDLAKGPGKDPGSRGITYASAGVDIEAGDRAIDLFKPLASKAT
                     RPEVRGGLGGFAGLFTLRGDYREPVLAASSDGVGTKLAIAQAMDKHDTVGLDLVAMVV
                     DDLVVCGAEPLFLLDYIAVGRIVPERLSAIVAGIADGCMRAGCALLGGETAEHPGLIE
                     PDHYDISATGVGVVEADNVLGPDRVKPGDVIIAMGSSGLHSNGYSLVRKVLLEIDRMN
                     LAGHVEEFGRTLGEELLEPTRIYAKDCLALAAETRVRTFCHVTGGGLAGNLQRVIPHG
                     LIAEVDRGTWTPAPVFTMIAQRGRVRRTEMEKTFNMGVGMIAVVAPEDTTRALAVLTA
                     RHLDCWVLGTVCKGGKQGPRAKLVGQHPRF"
     gene            complement(1211. .1393)
                     /gene="Rv0810c"
     CDS             complement(1211. .1393)
                     /gene="Rv0810c"
                     /note="Rv0810c, (MTV043.02c), len: 60. Unknown. TBparse
                     score is 0.876"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein Rv0810c"
                     /protein_id="CAA17616.1"
                     /db_xref="GI:2916868"
                     /db_xref="SPTREMBL:O53824"
                     /translation="MGRGRAKAKQTKVARELKYSSPQTDFQRLQRELSGTGTDRLDGD
                     GPSDDDSWNDEDDWRR"
     RBS             complement(1401. .1406)
                     /note="possible RBS for Rv0810c"
     gene            complement(1540. .2646)
                     /gene="Rv0811c"
     CDS             complement(1540. .2646)
                     /gene="Rv0811c"
                     /note="Rv0811c, (MTV043.03c), len: 368. Unknown but
                     equivalent to Mycobacterium leprae hypothetical protein
                     U2266F (366 aa). FASTA scores: U15182|MLU15182_13  (366
                     aa) opt: 1870 z-score: 2265.5 E(): 0;    77.4% identity in
                     367 aa overlap. TBparse score is 0.912"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein Rv0811c"
                     /protein_id="CAA17617.1"
                     /db_xref="GI:2916869"
```

```
                     /db_xref="SPTREMBL:O53825"
                     /translation="MAAVPAPDPGPDAGAIWHYGDPLGEQRAGQADAVLVDRSHRAVL
                     TLDGGDRQTWLHSISTQHVSDLPEGASTQNLSLDGQGRVEDHWIQTELGGTTYLDTEP
                     WRGEPLLAYLRKMVFWSMVTPRAADMAVLSLLGPRLAEERVLDALGLDVLPAEWLAVP
                     LAGGGIVRRMPDGLAGQIELDVVVKRGDRADWQRRLTQAGVRPAGIWAYEAHRVAHRV
                     PARRPRLGVDTDERTIPHEVGWIGGPGAGAVHLNKGCYRGQETVARVHNLGRPPRMLV
                     LLHLDESVQRPSTGDAVLAGGRTVGRLGTVVEHVELGPVALALLKRGLPGDTALVTGP
                     EAEVAAVIDVDSLPPADDVGAGRRAVERLRGGIR"
RBS                  2716. .2719
                     /note="possible RBS for Rv0812"
gene                 2729. .3598
                     /gene="pabC"
CDS                  2729. .3598
                     /gene="pabC"
                     /note="Rv0812, (MTV043.04), len: 289. pabC, similar to
                     sp|P28305|PABC_ECOLI 4-AMINO-4-DEOXYCHORISMATE LYASE (
                     (269 aa). FASTA scores: opt:  165 z-score: 205.5 E():
                     0.00064, (26.8% identity in 198 aa overlap). Also similar
                     to several ILVE proteins. TBparse score is 0.915"
                     /codon_start=1
                     /transl_table=11
                     /product="pabC"
                     /protein_id="CAA17618.1"
                     /db_xref="GI:2916870"
                     /db_xref="SPTREMBL:O53826"
                     /translation="MVVTLDGEILQPGMPLLHADDLAAVRGDGVFETLLVRDGRACLV
                     EAHLQRLTQSARLMDLPEPDLPRWRRAVEVATQRWVASTADEGALRLIYSRGREGGSA
                     PTAYVMVSPVPARVIGARRDGVSAITLDRGLPADGGDAMPWLIASAKTLSYAVNMAVL
                     RHAARQGAGDVIFVSTDGYVLEGPRSTVVIATDGDQGGGNPCLLTPPPWYPILRGTTQ
                     QALFEVARAKGYDCDYRALRVADLFDSQGIWLVSSMTLAARVHTLDGRRLPRTPIAEV
                     FAELVDAAIVSDR"
gene                 complement(3644. .4324)
                     /gene="Rv0813c"
CDS                  complement(3644. .4324)
                     /gene="Rv0813c"
                     /note="Rv0813c, (MTV043.05c), len: 226. Unknown but
                     highlysimilar to M. leprae cosmid
                     B226gp|U15182|MLU15182_16 (242aa). FASTA scores: opt: 1191
                     z-score: 1405.0 E(): 0; 78.3% identity in 226 aa overlap.
                     Also similar to mtcy05a6_37, Hypothetical protein, 30.4%
                     identity in 171 aa overlap."
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein Rv0813c"
                     /protein_id="CAA17619.1"
                     /db_xref="GI:2916871"
                     /db_xref="SPTREMBL:O53827"
                     /translation="MSSGAGSDATGAGGVHAAGSGDRAVAAAVERAKATAARNIPAFD
                     DLPVPADTANLREGADLNNALLALLPLVGVWRGEGEGRGPDGDYRFGQQIVVSHDGGD
                     YLNWESRSWRLTATGDYQEPGLREAGFWRFVADPYDPSESQAIELLLAHSAGYVELFY
                     GRPRTQSSWELVTDALARSRSGVLVGGAKRLYGIVEGGDLAYVEERVDADGGLVPHLS
                     ARLSRFVG"
misc_feature         complement(4487. .5652)
                     /note="region identical to positions (-1) 26726-27891 of
                     MTCY164, appropriate BAC clone coverage exists (Rv 419,
                     Rv134, Rv45)"
gene                 complement(4487. .4789)
                     /gene="sseC2"
CDS                  complement(4487. .4789)
                     /gene="sseC2"
                     /note="Rv0814c, (MTV043.06c), len: 100. Unknown but
                     identical to Mycobacterium tuberculosis MTCY164.28 (100
                     aa). Equivalent to hypothetical M. leprae protein G699208
                     U2266J (143 aa) and similar to S.erythraea rhodanese-like
                     protein M29612|SERCYSA_5 (101 aa). FASTA scores:
                     Z95150|MTCY164_29 (100 aa)  opt:  647 z-score: 865.3 E():
                     0;   100.0% identity in 100 aa overlap; U15182|MLU15182_17
                     (143 aa)  opt:  545 z-score: 729.4 E(): 0;    84.0%
                     identity in 100 aa overlap; and M29612|SERCYSA_5  (101 aa)
                     opt:  345 z-score: 469.6 E():1.2e-18;    57.1% identity in
                     98 aa overlap. TBparse scoreis 0.853"
                     /codon_start=1
                     /transl_table=11
                     /product="sseC2"
                     /protein_id="CAA17620.1"
                     /db_xref="GI:2916872"

  Query Match           46.7%;  Score 417.6;  DB 1;  Length 68848;
  Best Local Similarity  68.9%;  Pred. No. 5e-86;
  Matches  573;  Conservative    0;  Mismatches  259;  Indels    0;  Gaps    0;

Qy     42 GCCAAGTTCTCTGTCAAGCAATAATAATCTCAGAAAATACAACAGGAGTATCTAAAGCGA 101
Db   8976 GCACGGATAACTGACCGAAGATGTTCCCCCCCAAGAAGTCGTAAGCAGGAGCTTAAGTGG 9035

Qy    102 TGTCGAAGCTCAAGCTCAATGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGA 161
Db   9036 CCAAGCGGTTGGACCTCACGGACGTCAACATCTACTACGGGTCATTTCATGCGGTCGCTG 9095

Qy    162 ACGTGAACCTCGAGGTTCCTGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTG 221
Db   9096 ATGTGTCGCTGGCGATTCTGCCCCGCAGCGTCACGGCGTTCATCGGTCCCTCGGGCTGCG 9155

Qy    222 GCAAGTCCACAGTTCTCCGCTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACG 281
Db   9156 GCAAGACGACGGTGCTGCGCACCTTGAACCGGATGCATGAGGTCATCCCCGGAGCTCGAG 9215

Qy    282 TCAAGGGCGAGATCCTTCTCGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTG 341
Db   9216 TCGAGGGTGCCGTACTGCTCGATGATCAAGATATCTACGCCCCCGGTATCGACCCGGTCG 9275

Qy    342 CAGTCCGTAACACCATCGGCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCA 401
Db   9276 GTGTCCGCCGGGCAATCGGGATGGTGTTTCAGCGGCCGAATCCATTCCCCGCCATGTCGA 9335

Qy    402 TCGAGGACAACGTGGTTGCAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCA 461
Db   9336 TTCGCAACAATGTGGTTGCCGGCCTGAAGCTGCAGGGTGTGCGCAATCGCAAGGTGCTCG 9395

Qy    462 AGGAAGTTGCTGAGAAGTCTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTC 521
Db   9396 ACGATACGGCCGAATCCTCGCTGCGCGGCGCAAACCTGTGGGACGAGGTCAAGGATCGAC 9455

Qy    522 TGGACAAGCCAGGCGGCGGCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCG 581
Db   9456 TGGATAAACCCGGCGGCGGATTGTCTGGGGGGCAGCAGCAGCGGTTGTGCATCGCACGGG 9515

Qy    582 CGATCGCGGTTGAGCCAGAAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAA 641
Db   9516 CAATCGCCGTGCAACCCGACGTGTTGCTGATGGACGAGCCCTGCTCCTCGCTGGACCCAA 9575

Qy    642 TCTCAACCCTGGCTGTGGAGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCA 701
Db   9576 TCTCCACCATGGCCATCGAAGACCTGATCAGCGAGCTCAAGCAGCAGTACACCATCGTCA 9635

Qy    702 TCGTGACCCACAACATGCAGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCC 761
Db   9636 TCGTCACCCATAACATGCAGCAGGCTGCCCGGGTGAGTGATCAGACGGCATTCTTCAACC 9695

Qy    762 TGGAGGCGACCGGTAGGCCAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCG 821
Db   9696 TGGAAGCGGTGGGAAAGCCGGGGCGGCTGGTAGAGATCGCCAGCACCGAGAAAATCTTCT 9755

Qy    822 AAAACCCAGATCAGAAGGAAACCGAAGATTACATCTCCGGCCGCTTCGGATA 873
Db   9756 CCAACCCGAACCAGAAGGCCACCGAGGACTACATCTCCGGGCGCTTCGGCTA 9807

RESULT 8
SCD84/c
LOCUS       SCD84                   32354 bp    DNA     linear   BCT 12-MAY-2002
DEFINITION  Streptomyces coelicolor cosmid D84.
ACCESSION   AL353816 AL645882
VERSION     AL353816.1  GI:7636006
KEYWORDS    ABC transport system ATP-binding protein; ABC transport system
            transmembrane protein; acetyltransferase; ECF subfamily sigma
```

factor; glnR; hydrolase; integral membrane protein; LacI-family regulatory protein; lipoprotein; molybdopterin converting factor; mutT-like protein; phosphate ABC transport system ATP-binding protein; phosphate ABC transport system permease protein; phosphate transport protein; phosphate-binding protein precursor; polyphosphate kinase; protease; secreted 5'-nucleotidase; secreted protein; thiosulfate sulfurtransferase; two-component system sensor; two-component systen response regulator.

SOURCE Streptomyces coelicolor A3(2).
ORGANISM Streptomyces coelicolor A3(2)
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE 1 (bases 1 to 32354)
AUTHORS Redenbach,M., Kieser,H.M., Denapaite,D., Eichner,A., Cullum,J., Kinashi,H. and Hopwood,D.A.
TITLE A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome
JOURNAL Mol. Microbiol. 21 (1), 77-96 (1996)
MEDLINE 97000351
PUBMED 8843436

REFERENCE 2 (bases 1 to 32354)
AUTHORS Saunders,D.C. and Harris,D.
JOURNAL Unpublished

REFERENCE 3 (bases 1 to 32354)
AUTHORS Cerdeno,A.M., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
TITLE Direct Submission
JOURNAL Submitted (20-APR-2000) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK

COMMENT Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.
(URL; http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg SC7B7.01c. SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand).
The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database.
The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at http://www.nih.go.jp/
jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.
IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid D84.

FEATURES Location/Qualifiers

source 1. .32354
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"
/db_xref="taxon:100226"
/clone="cosmid D84"

misc_feature 1. .111
/note="Nominal overlap with Streptomyces coelicolor cosmid D84
nominal overlap with Streptomyces coelicolor cosmid StD72A"

misc_feature complement(417. .449)
/note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"

RBS complement(487. .491)

gene 1083. .1976
/gene="SCD84.02"
/note="SCO4135"

CDS 1083. .1976
/gene="SCD84.02"
/note="SCD84.02, possible lipoprotein, len: 297 aa; similar to TR:CAB75314 (EMBL:AL139164) Streptomyces coelicolor putative lipoprotein SCC46.10, 285 aa; fasta scores: opt: 180 z-score: 203.0 E(): 0.00075; 22.4% identity in 232 aa overlap. Contains correctly situated match to Prosite entry PS00013 Prokaryotic membrane lipoprotein lipid attachment site and possible N-terminal region signal peptide sequence".
/codon_start=1
/transl_table=11
/product="putative lipoprotein"
/protein_id="CAB88468.1"
/db_xref="GI:7636008"
/db_xref="SPTREMBL:Q9KZW6"
/translation="MVGRERELRSGWAGGGPALAGGAVCVCLALAGCSGSGQGASVQG AGGGDPVAVLHRAADRLVAAGSAKARTAMQMASGGTRVTIRGEGVYDFEERLGQLKVR LPQDPAGTSEHRPITELLAPGALFMKNRGAGVPDDKWVRVETDSLSDGNLVTGGATDP FAAAEVLRGTRTATFVGRTELAGTEVRHYRGTADLARAARGASAGNRKALAAAAQGFA TAQVPFDAYLDDQGRIRKVRHRFSFANGQEDGTGIVDVASTTLLYDFGAAVRVRLPED QDIYAGKIAAEPGERGVDGKN"

misc_feature 1149. .1181
/gene="SCD84.02"
/note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"

RBS 2042. .2045

gene complement(2204. .2605)
/gene="SCD84.03c"
/note="SCO4136"

CDS complement(2204. .2605)
/gene="SCD84.03c"
/note="SCD84.03c, hypothetical protein, len: 133 aa; similar to TR:O06070 (EMBL:Z95398) Mycobacterium leprae hypothetical 15.2 kD protein MLCL622.07, 135 aa; fasta scores: opt: 374 z-score: 470.4 E(): 9.6e-19; 55.9% identity in 102 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein SCD84.04c"
/protein_id="CAB88469.1"
/db_xref="GI:7636009"
/db_xref="SPTREMBL:Q9KZW5"
/translation="MTTTEAGASAPSPAVDGAVNQTVRQAETDGTDIVTDHDRGVHGY HKQKAEHLKRLRRIEGQIRGLQRMVDEDVYCIDILTQVSASTKALQSFALQLLEEHLR HCVADAALKGGTEIDAKVEEATKAIGRLLRT"

RBS complement(2611. .2615)

RBS 2802. .2806

gene 2813. .3433
/gene="SCD84.04"
/note="SCO4137"

CDS 2813. .3433
/gene="SCD84.04"
/note="SCD84.04, hypothetical protein, len: 206 aa; similar to TR:AAF10501 (EMBL:AE001945) Deinococcus radiodurans conserved hypothetical protein DR0924, 212 aa; fasta scores: opt: 347 z-score: 407.5 E(): 3.1e-15; 32.4% identity in 210 aa overlap. Contains Pfam match to entry PF01865 DUF47, Protein of unknown function"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein SCD84.04"
/protein_id="CAB88470.1"
/db_xref="GI:7636010"
/db_xref="SPTREMBL:Q9KZW4"
/translation="MRFRLTPRETSFYDMFAASADNIVTGSKLLMELLGADASARAEI

```
                    AERMRAAEHAGDDATHAIFHQLNSSFITPFDREDIYNLASSLDDIMDFMEEAVDLVVL
                    YNVEELPKGVEQQIEVLARAAELTAEAMPNLRTMDNLTEYWIEVNRLENQADQIHRKL
                    LAHLFNGKYEAIEVLKLKQIVDVLEEAADAFEHVANTVETIAVKES"
     misc_feature   2813. .3424
                    /gene="SCD84.04"
                    /note="Pfam match to entry PF01865 DUF47, Protein of
                    unknown function, score 66.30, E-value 6.7e-16"
     gene           3441. .4439
                    /gene="SCD84.05"
                    /note="SCO4138; pitH"
     CDS            3441. .4439
                    /gene="SCD84.05"
                    /note="SCD84.05, pitH, phosphate transport protein, len:
                    332 aa; similar to N-terminal region of SW:PITH_STRHA
                    (EMBL:L05390) Streptomyces halstedii putative low-affinity
                    inorganic phosphate transporter (fragment), 213 aa; fasta
                    scores: opt: 768 z-score: 850.1 E(): 0; 59.8% identity in
                    189 aa overlap, to TR:O30499 (EMBL:AF008187) Sinorhizobium
                    meliloti phosphate transport protein, Pit, 334 aa; fasta
                    scores: opt: 988 z-score: 1087.8 E(): 0; 48.3% identity in
                    327 aa overlap and to TR:CAB59459 (EMBL:AL132644)
                    Streptomyces coelicolor putative secreted protein SCI8.28,
                    423 aa; fasta scores: opt: 991 z-score: 986.9 E(): 0;
                    45.9% identity in 331 aa overlap. Contains Pfam match to
                    entry PF01384 PHO4, Phosphate transporter family. Contains
                    also possible hydrophobic membrane spanning regions"
                    /codon_start=1
                    /transl_table=11
                    /product="phosphate transport protein"
                    /protein_id="CAB88471.1"
                    /db_xref="GI:7636011"
                    /db_xref="SPTREMBL:Q9KZW3"
                    /translation="MDTFALVVTIGVALFFTYTNGFHDSANAIATSVSTRALTPRAAL
                    AMAAVMNLAGAFMGSGVAKTVSEGVIETPEGSKGMGILFAALVGAIVWNLITWYFGLP
                    SSSSHALFGGMVGAALAGSTTVYWHGVVDKIVIPMFVSPVIGLLAGYLVMTAIMWIFR
                    RANPHKAKRGFRIAQTVSAAGMALGHGLQDAQKTMGIVVMALVIADVEDYGDPIPVWV
                    KLACALMLSLGTYAGGWRIMRTLGRKIIELDPPQGFAAETTGASIMFGSAFIFHAPIS
                    TTHVITSAIMGVGATRRVNAVRWGVAKNIIMGWFITMPAAAVVAAVSFWIVNLAVL"
     misc_feature   3498. .4400
                    /gene="SCD84.05"
                    /note="Pfam match to entry PF01384 PHO4, Phosphate
                    transporter family, score 140.70, E-value 2.5e-38"
     gene           complement(4559. .5335)
                    /gene="SCD84.06c"
                    /note="SCO4139; pstB"
     CDS            complement(4559. .5335)
                    /gene="SCD84.06c"
                    /note="SCD84.06c, pstB, phosphate ABC transport system
```

Query Match 46.6%; Score 416.4; DB 1; Length 32354;
Best Local Similarity 71.7%; Pred. No. 8.7e-86;
Matches 546; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

```
Qy    117 TCAATGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGG 176
Db   5316 TCAGCGGCCTCAACGCCTACTACGGCTCCTTCCTGGCCATCGAGGACATCTCGATGACCG 5257

Qy    177 TTCCTGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTC 236
Db   5256 TCGAGCCCCGCTCCGTGACGGCCTTCATCGGCCCCTCCGGCTGCGGCAAGTCCACCTTCC 5197

Qy    237 TCCGCTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCC 296
Db   5196 TGCGCACGCTCAACCGCATGCACGAGGTCACCCCGGGCGGCCGCGTCGACGGCAAGGTGA 5137

Qy    297 TTCTCGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCA 356
Db   5136 TGCTCGACGACGAGAACCTCTACGGCGCGGGCATCGACCCGGTGGCCGTGCGCCGCGAGG 5077

Qy    357 TCGGCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGG 416
Db   5076 TCGGCATGGTCTTCCAGCGCCCGAACCCGTTCCCCACGATGTCGGTGTACGACAACGTGG 5017
```

```
Qy    417 TTGCAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGA 476
Db   5016 CGGCGGGCCTGCGCCTGAACGGCAAGTACAAGAAGTCCCAGCTGGACGACGTCGTCGAGA 4957

Qy    477 AGTCTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCG 536
Db   4956 AGTCGCTGCGCGGCGCCAACCTCTGGAACGAGGTCAAGGACCGTCTGAACAAGCCGGGCT 4897

Qy    537 GCGGCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGC 596
Db   4896 CCGGCCTCTCCGGCGGCCAGCAGCAGCGCCTGTGCATCGCCCGCGCGATCGCGGTCGAGC 4837

Qy    597 CAGAAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTG 656
Db   4836 CCAACGTGCTCCTGATGGACGAGCCCTGCTCCGCCCTCGACCCGATCTCCACCCTCGCCA 4777

Qy    657 TGGAGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACA 716
Db   4776 TCGAGGACCTGATCGGCGAGCTGAAGGAGCGCTTCACGATCGTCATCGTGACGCACAACA 4717

Qy    717 TGCAGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTA 776
Db   4716 TGCAGCAGGCGGCGCGCGTCTCGGACCGTACGGCGTTCTTCAACCTGGCGGCGGTCGGCC 4657

Qy    777 GGCCAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGA 836
Db   4656 AGCCCGGCAAGCTGATCGAGATCGACGAGACGGAGCGCATCTTCTCCAACCCGTCGGTCC 4597

Qy    837 AGGAAACCGAAGATTACATCTCCGGCCGCTTCGGATAAATCG 878
Db   4596 AGGCCACGGAGGACTACATCTCCGGCCGCTTCGGCTAGACCG 4555
```

```
RESULT 9
SGR243674
LOCUS       SGR243674               9302 bp    DNA     linear   BCT 06-JAN-2001
DEFINITION  Streptomyces griseus subsp. griseus (partial) ppk, mutT, pstS,
            pstC, pstA, pstB, pitH genes and ORF.
ACCESSION   AJ243674
VERSION     AJ243674.3  GI:12054826
KEYWORDS    mutT gene; pitH gene; polyphosphate kinase; ppk gene; pstA gene;
            pstB gene; pstC gene; pstS gene.
SOURCE      Streptomyces griseus subsp. griseus.
  ORGANISM  Streptomyces griseus subsp. griseus
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 9302)
  AUTHORS   Siekmoeller,J., Fuechtenbusch,B., Wittchen,K.D., Kranz,C. and
            Pape,H.
  TITLE     Characterization of a membrane protein participating in the
            in-vitro formation of macrotetrolides from nonactinic acid and
            genes coding for a putative phosphate uptake system from
            Streptomyces griseus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 9302)
  AUTHORS   Pape,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-1999) Pape H., Institut fuer Mikrobiologie,
            Universitaet Muenster, Corrensstrasse 3, D-48149 Muenster, GERMANY
  REMARK    revised by [3]
REFERENCE   3  (bases 1 to 9302)
  AUTHORS   Pape,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2000) Pape H., Institut fuer Mikrobiologie,
            Universitaet Muenster, Corrensstrasse 3, D-48149 Muenster, GERMANY
COMMENT     On Jan 8, 2001 this sequence version replaced gi:6689250.
FEATURES             Location/Qualifiers
     source          1. .9302
                     /organism="Streptomyces griseus subsp. griseus"
                     /strain="ETH A7796"
                     /db_xref="taxon:67263"
     gene            1. .2114
                     /gene="ppk"
```

```
CDS             <1. .2114
                /gene="ppk"
                /codon_start=3
                /transl_table=11
                /product="polyphosphate kinase"
                /protein_id="CAC21099.1"
                /db_xref="GI:12054827"
                /translation="IRSGRDADAYEPDRDGDELPQGRFLDRERSWLAFNERVLELAED
                PATPILERANFLAIFASNLDEFFMVRVAGLKRRIATGVATRSASGLQPREVLDLIWTR
                SRELMARHAACYQQDIAPDLSDEGIQLIRWPDLTEKEQARLFTFFRQRVFPVLTPLAV
                DPAHPFPYISGLSLNLAVVVRNPVSGHRHFARVKVPPLLTRFLEASPQRYVPIEDIIA
                AHLEELFPGMEVLAHHMFRVTRNEDLEVEEDDAENLLQALEKELMRRRFGPPVRLEVE
                ESIDPYVLDLLVRELKVSDAEVYPLPGPLELTGLFAIASLDRPELKFPKFIAGTHRDL
                AEVESASAPDIFAALRERDVLLHHPYDSFSTSVQAFLEQAAGDPDVLAIKQTLYRTSG
                DSPIVDALIDAAESGKQVLVLVEIKARFDEQANIKWARKLEEAGCHVVYGLVGLKTHC
                KLSLVVRQEGDTLRRYSHVGTGNYHPKTARLYEDLGLLTGDPQVGADLSDLFNRLSGY
                SRRETYRRLLVAPKSLRDGLIARINKEIAHHRAGRPAYVRFKVNSMVDEAIIDACCRA
                AQAGVPVDIWVRGICAIRPASRLSENIRVRSILGRFLEHSRVFSFGNGGEPEVWFGSA
                DMMHRNLDRRIEALVRVTDPAHRAALSRLLETGMADTTASWHLGPDGNWTRHATDADG
                QPLRHVQEMLIEARRRRRATP"
gene            2224. .3237
                /gene="mutT"
CDS             2224. .3237
                /gene="mutT"
                /note="putative"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC21100.1"
                /db_xref="GI:12054828"
                /translation="MLARSPDPAGPRPAGDLTTEAALAPYLRGQAAEFLRSLRLHHEH
                SAPADSGGHDAAAATRALRRSARRISGSLHTFRAALDPLWADQLRAELAWLSGILARE
                HAYANRLTRLVEALHQLSGPALPAARGVRAAGESATSDAQGRAALGVGAARAGALLER
                RLTLARTRSHSAALQALGSSRFHAVADAVALLASDVPLAPGTAGRTAEALLEPAERAE
                QRLLTAVAALPPDDTEPYNEAQDAAWHQARLLLRLHRYAHEVVLGAADPALTGAGHAL
                DLHRDAVEAAGAAAAAARTPRIAPATAYALGVLHADQRHEVEAARTVFRETWPYAAAL
                STP"
CDS             3234. .3692
                /note="ORF"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="CAC21102.1"
                /db_xref="GI:12054829"
                /translation="MSGGPHRAPGEHPTRTGTVLAAGCVLWRRAPDGDGDDVEVCLVH
                RPRYDDWSFPKGKLKRGEEPLAAAVREVLEETGHHCAPGDRLPTSRYAVDGRPKEVAY
                WAAEATGGSFTANDEVDRILWLTPDAARVRLTRPRDVRQLIALLDALAAS"
gene            3891. .5027
                /gene="pstS"
CDS             3891. .5027
                /gene="pstS"
                /codon_start=1
                /transl_table=11
                /product="PstS protein"
                /protein_id="CAB65418.2"
                /db_xref="GI:12054830"
                /translation="MKLQRKNRLRATALGALAVSGALVLTACGSDDNTGGGTAGTGEK
                TTAASNVDCGKAKGQLRASGSSAQKNAMDLWVKNYMAACSGVEINYNSSSSGEGIVAF
                NQGTVGFAGSDSALKPEEVADSKKMCKTGQGINLPMVGGPVAIGFHLEGVDALTLDAP
                TLAKIFDSKIKKWNDPAIAKLNDGVKLPDKAIQAFHRSEDSGTTQNLGKYLGAAAPSD
                WKYEAEKKWPAPGGQAASGSSGVAAQVKQVDGAIGYFELSYAKSQSIPTVDINTGGAT
                PVKATSENASKAIAAAKIKGTGKDLALDLDYTTKAEGAYPLVLVTYEVVCDTGNKPET
                LDTVKSFLSYTASDDGQKILTDAGYAPIPAEINAKVRETIGSLS"
sig_peptide     3891. .3971
                /gene="pstS"
gene            5220. .6206
                /gene="pstC"
CDS             5220. .6206
                /gene="pstC"
                /codon_start=1
                /transl_table=11
                /product="PstC protein"
                /protein_id="CAB65970.1"
                /db_xref="GI:12054831"
                /translation="MASTTPIDTPPGRRPAAQPPVHRRMGDKIFLGLSRGSGILLLVI
                MASIAVFLSYRSAIAISENEGNFLTTFDWNPAGDPPVFGIAVLLFGTVVSSIIAMVIA
                VPIAVGIALFISHYAPRKLAAPIAYVVDLLAAVPSIVYGIWGALVLVPYLEGLNLWLD
                DFFGWTYLFDKTEIGIPRSLFTVGILLAIMILPIVTSVTREVFLQVPKMNEEAALALG
                ATRWEVIRLSVLPFGRSGIISASMLGLGRALGETMAVATVLSPSFLISLHVLNPGGGT
                FAQNIAAKFARPTRSGGTLIASGLVLFVLTLLVNGAARLIIARRKEYSGANA"
gene            6203. .7315
                /gene="pstA"
CDS             6203. .7315
                /gene="pstA"
                /codon_start=1
                /transl_table=11
                /product="PstA protein"
                /protein_id="CAC21104.1"
                /db_xref="GI:12054832"
                /translation="MSHAPTGVIERPTPASPTPKSSLGSRSLPRLAPLGFAVVSVALG
                IGIPLAAGWHRRIQWGMLSALFFLAISYVVTTIVENQRQARDRLATSVMWVCFLIAVV
                PLASLLWTTIVRGAERLDGYFLTHSMAGVLGSEHSGGVYHALIGTLEQVGIATVISVP
                FGLLTAVYLVEYGKGTLAKAVTFFVDVMTGIPSIVAGLFILSIMLIASIEPSGLMGAL
                ALAILMVPVVVRSTEEMLKLVPNELREASLALGIPKWRTILKVVLPTAIGGITTGVML
                AIARIAGETAPIILLVFGSQLINTNPFEAPSPTAVLHLRAVPDRRGRFLRPRLAAALV
                LIAFVMILNLVARGIARWKAPKTGADAATATSRKKQ"
gene            7321. .8097
                /gene="pstB"
CDS             7321. .8097
                /gene="pstB"
                /codon_start=1
                /transl_table=11
                /product="PstB protein"
                /protein_id="CAC21105.1"
                /db_xref="GI:12054833"
                /translation="MAKRIDISGLSAFYGNHKAIEDISMTVEPRSVTAFIGPSGCGKS
                TFLRTLNRMHEVTPGGRVEGKVLLDDENLYDSNVDPVTVRRTVGMVFQRPNPFPTMSI
                FDNVAAGLRLNGSYRKSELNDVVEKSLRGANLWNEVKDRLNKPGSGLSGGQQQRLCIA
                RAIAVEPQVLLMDEPCSALDPISTLAIEDLIGELKERFTIVIVTHNMQQAARVSDRTA
                FFNLAAVGKPGRLIEIDETERIFSNPSVQATEDYISGRFG"
gene            complement(8243. .9241)
                /gene="pitH"
CDS             complement(8243. .9241)
                /gene="pitH"
                /codon_start=1
                /transl_table=11
                /product="PitH protein"
                /protein_id="CAC21106.1"
                /db_xref="GI:12054834"
                /translation="MDTFALIVTIGVALGFTYTNGFHDSAQRIATSVSTRALTPRAAL
                AMAAVMNLAGAFLGQGVAKTVSEGLIATPVGQKGMGILFAALVGAIIWNLITWYYGLP
                SSSSHALFGGMVGAALAGGTDVIWSGVLKKVVIPMFISPFVGLIAGYLVMVGIMWIFR
                KSNPHKAKRGFRIAQTVSAAGMALGHGLQDAQKTMGIVVMALVIADVEGPNDEIPVWV
                KIACALMLSLGTYAGGWRIMRTLGRKIIELDPPQAFAAETTGASIMFGSAFLFHAPIS
                TTHVITSAIMGVGATKRVNAVRWGVAKNIILGWFITMPAAALVAALSYGAVLLLFG"
BASE COUNT      1431 a   3657 c   2910 g   1304 t
ORIGIN
```

```
  Query Match           46.1%;  Score 412;  DB 1;  Length 9302;
  Best Local Similarity  71.2%;  Pred. No. 7.8e-85;
  Matches  544;  Conservative    0;  Mismatches  220;  Indels    0;  Gaps    0;

Qy     111 TCAAGCTCAATGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACC 170
           || |  |||  |   ||| |  || ||||||  |   | | ||  |  || || |
Db    7334 TCGACATCAGCGGCCTCAGCGCCTTCTACGGCAACCACAAGGCCATCGAGGACATCTCGA 7393

Qy     171 TCGAGGTTCCTGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCA 230
           |    ||     | ||||| || ||||| |||||||| || |||||||| ||||||||||
Db    7394 TGACCGTGGAGCCCCGCTCCGTGACCGCCTTCATCGGCCCGTCCGGCTGCGGCAAGTCCA 7453

Qy     231 CAGTTCTCCGCTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCG 290
           |  | || ||| || | || || ||||||||||| ||||| || |    |||| |||||
Db    7454 CCTTCCTGCGCACCCTGAACCGGATGCACGAGGTGACCCCCGGCGGCCGCGTCGAGGGCA 7513

Qy     291 AGATCCTTCTCGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTA 350
           || | || || |||| ||||||| |||||| ||||||  ||||||| ||  | || ||
Db    7514 AGGTGCTGCTGGACGACGAGAACCTCTACGACTCCAACGTCGACCCGGTCACCGTGCGCC 7573
```

```
Qy     351 ACACCATCGGCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACA 410
           |||  | ||||||||||||||||    | ||||| ||||| |||||||| |||   ||||
Db    7574 GCACGGTGGGCATGGTCTTCCAGCGCCCGAACCCCTTCCCCACCATGTCGATCTTCGACA 7633

Qy     411 ACGTGGTTGCAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTG 470
           ||||||  || || |||  |||   ||||    |   |||||   |||| || || || |
Db    7634 ACGTGGCGGCGGGCCTGCGGCTGAACGGCAGCTACCGCAAGAGCGAGCTGAACGACGTCG 7693

Qy     471 CTGAGAAGTCTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGC 530
           ||||||||  || || ||||| ||||| ||| | ||||| ||||| || ||  ||||||
Db    7694 TGGAGAAGTCCCTGCGCGGCGCGAACCTCTGGAACGAGGTCAAGGACCGCCTCAACAAGC 7753

Qy     531 CAGGCGGCGGCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGG 590
           | |||  |||||||||||| || |||||||||| |||||||||| ||||| |||||||
Db    7754 CCGGCTCCGGCCTCTCCGGCGGCCAGCAGCAGCGCCTGTGCATCGCCCGCGCCATCGCGG 7813

Qy     591 TTGAGCCAGAAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCC 650
           | |||||  |  |||| || |||||||||| ||||| ||||| ||||| ||||| || |
Db    7814 TCGAGCCGCAGGTCCTGCTGATGGACGAGCCGTGCTCGGCGCTGGACCCGATCTCGACGC 7873

Qy     651 TGGCTGTGGAGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCC 710
           | ||  | ||||||| |||  |||||||||||    ||||| |||||||||||||| |
Db    7874 TCGCCATCGAGGACCTGATCGGCGAGCTGAAGGAGCGCTTCACGATCGTCATCGTGACGC 7933

Qy     711 ACAACATGCAGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGA 770
           |||||||||||||||| || || || || || |  || ||||||| |  ||| |  |||
Db    7934 ACAACATGCAGCAGGCGGCCCGCGTCTCGGACCGCACGGCGTTCTTCAACCTCGCCGCGG 7993

Qy     771 CCGGTAGGCCAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAG 830
           ||| | ||| || || ||  | || || |     ||  ||  |||||||   |||||
Db    7994 TCGGCAAGCCCGGCCGGCTCATCGAGATAGACGAGACGGAGCGGATCTTCTCCAACCCGT 8053

Qy     831 ATCAGAAGGAAACCGAAGATTACATCTCCGGCCGCTTCGGATAA 874
           |||  || || || ||||||||||||||||||||| | |
Db    8054 CGGTCCAGGCCACGGAGGACTACATCTCCGGCCGCTTCGGCTGA 8097
```

RESULT 10
MLU15182

LOCUS	MLU15182	40123 bp DNA linear BCT 11-MAR-2002
DEFINITION	Mycobacterium leprae cosmid B2266.	
ACCESSION	U15182	
VERSION	U15182.1 GI:699197	
KEYWORDS	.	
SOURCE	Mycobacterium leprae.	
ORGANISM	Mycobacterium leprae Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.	
REFERENCE	1 (bases 1 to 40123)	
AUTHORS	Smith,D.R.	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 40123)	
AUTHORS	Robison,K.	
TITLE	Direct Submission	
JOURNAL	Submitted (28-SEP-1994) Department of Genetics, Harvard Medical School, 200 Longwood Avenue, Boston MA 02115	
COMMENT	This sequence data was produced by the Genome Sequencing Center located at Genome Therapeutics Corp. (formerly Collaborative Research Corp.) (100 Beaver St., Waltham MA, 02154 617-893-5007). Please contact Doug Smith (smith@cr.cric.com) for further information. The annotation should be considered preliminary and incomplete.	

```
FEATURES             Location/Qualifiers
     source          1. .40123
                     /organism="Mycobacterium leprae"
                     /db_xref="taxon:1769"
     CDS             complement(81. .119)
                     /codon_start=1
                     /transl_table=11
                     /product="u2266a"
                     /protein_id="AAA62981.1"
                     /db_xref="GI:699217"
                     /translation="MGVPVSHQYNNR"
     CDS             complement(309. .392)
                     /codon_start=1
                     /transl_table=11
                     /product="u2266b"
                     /protein_id="AAA62980.1"
                     /db_xref="GI:699216"
                     /translation="MLFDVRADTMDVCLELACHVTESQWQV"
     CDS             complement(332. .772)
                     /codon_start=1
                     /transl_table=11
                     /product="u2266c"
                     /protein_id="AAA62979.1"
                     /db_xref="GI:699215"
                     /translation="MNSQRSAALAGMPNIMAVGDDVVDARVGIVTPFQLQSILAPVTS
                     TAIFLGVTIDISGEATVHDTLSERSGLICVIASRSYQPLITGDSESVLTPGIDCLAAL
                     GPQNCIHSSRTGQTTAHLLSEHTETCCLTFVPIRWTYVWSWPAT"
     CDS             687. .2018
                     /codon_start=1
                     /transl_table=11
                     /product="pepX"
                     /protein_id="AAA62998.1"
                     /db_xref="GI:699234"
                     /translation="MPTLASTTSSPTAMMLGMPASAADLCEFINASPSPFHVCATVAG
                     RLLDAGYAELSEVERWPDHPGRYFIVRAGSLVAWSAGQGVKAHAPFRIVGAHTDSPNL
                     RVKQHPDLLVAGWRVVALQPYGGAWLNSWLDRDLGVSGRLSVRSAGKGSEITDRLVRI
                     DDPILRVPQLAIHLAEDRKSLTLDPQRHVNAVWGVGDKAGSLLEYVAERTGVAVADVL
                     AVDLMTHDLVPSMVIGADANLLSAPRLDNQVSCYAGMEALLASVPHDCLPVLALFDHE
                     EVGSTSDRGARSNLLSTVLERIVLAAGGGRDDYLRRLPASLLVSADMAHATHPNYPEC
                     HEPSHLIEVNAGPVLKVHPNLRYATDGRTAAAFEVACQQAGVRLQRYEHRADRPCGST
                     IGPLASARTGIPTVDVGAAQLAMHSARELMGAHDVAVYSAALQAFFSADLF"
     CDS             2870. .5134
                     /codon_start=1
                     /transl_table=11
                     /product="purL"
                     /protein_id="AAA62986.1"
                     /db_xref="GI:699222"
                     /translation="MIDTVEYAATTPDQPQPFAELGLREDEYQRVREILGRRPTDTEL
                     AMYSVMWSEHCSYKSSKVHLRYFGETTTEEMRTGMLAGIGENAGVVDIGDGWAVTFKV
                     ESHNHPSYVEPYQGAATGVGGIVRDIMAMGARPVAVMDQLRFGAADALDTRRVLDGVV
                     RGIGGYGNSLGLPNIGGETVFDSCYDGNPLVNALCVGVLRQEDLHLAFASGAGNKIIL
                     FGACTGLDGIGGVSVLASDTFDAEGARKKLPSVQVGDPFMEKVLIECCLELYAGGLVI
                     GIQDLGGAGLSCATSELASAGDVGMAIQLDTVPRRAKDMTPAEVFCSESQERMCAVVA
                     PENVDAFLAVCRKWEVLATVIGEVTDGDRLRITWHGETVVDVPPRTVAHEGPVYQRPV
                     SRPESQEALNADSSKGLPRPVSGDELRATLLALLGSPHLCSRAFITEQYDRYVRGNTV
                     LAEHADAGVLRIDESTGRGIALSTDASGRYTRLDPYAGAQLALAEAYRNVAVTGATPV
                     AVTNCLNFGSPEDPGVMWQFAQAVRGLADGCAALKIPVTGGNVSFYNQTGAVAILPTP
                     VVGVLGVLDNVARRIHTSLGTEPGEILMLLGDTYDEFDGSVWAQVMAGHLGGLPPMVD
                     LAREKLLAEVLSSASRDELVSAAHDLSEGGLAQAIVESALAGETGCRIALPEDADPFV
                     MLFSESAGRVLVAVPRPEESRFRSMCEARGLPAMRIGVVDQGSDSIEVRGQFTVSLAE
                     LRMTFEAVLPRFFG"
     tRNA            complement(3233. .3315)
                     /product="tRNA-Glu"
     CDS             5267. .6220
                     /codon_start=1
                     /transl_table=11
                     /product="icc"
                     /protein_id="AAA62995.1"
                     /db_xref="GI:699231"
                     /translation="MHRLRAAEHPRPDYVLLHISDTHLISDGSLYGAVDADSRLGELL
                     EQLKHSQLRPDAIIFTGDLADRGEPEAYRKLRCLVESFATELDAELFWVMGNHDNRVA
                     LRTLLLDEAPSMAPLDGVRRVDGLRVITLDTSVPGRHYGEISASQLDWLADELTTSAP
                     DGTILALHHPPIPSVLDMAVTVELRDQASLGRVLKGSDIRAILAGHLHYSTNATFVGI
                     PVSVASATCYTQDLTVVAGGARGRDGAQGCNLVHVYQDTVVHSVIPLGIGNMVGTFVS
                     PAQARRDIAESGLFIEPSGRDSLFAHPPMVLTPSVTQSLGD"
     CDS             complement(6135. .7763)
                     /codon_start=1
                     /transl_table=11
                     /product="cpsY"
                     /protein_id="AAA62996.1"
                     /db_xref="GI:699232"
                     /translation="MSKIVSCEDDRPVRRTLEPIIVTRQGKVARLESSLTPHEAQIED
```

```
                LIFLRKALNRADIPFLFIRNHKNRPVLAINIKLRPAVERALVTACASEPMYAKTIDER
                GLSPVLVAKGQLSQSIDPRIVRLYRRRIAPGGFRFGSRFGVELQFWSFEETLIRCPVE
                NSLTRKVLPRKEVTPATIKLYGYKWHTIEGMFTPHASDVTFDIDLVFSWVDGSDPEFR
                ARRAAEMSHHVVGEGDDADARIRQIDELKYALRSVNMFAPWIRRIFIATDSIPPSWLA
                DHPMITIVPAEDHFSDRSALPTYNSHAVESQLHRIPDLSEHFLYSNDDMFFGRPLKAS
                MFFSPGGVTRFIEAKTRIGLGTNDPTRSGFENAARVNRQLLLRRFGQLITRHLEHTTV
                PLRKSVLFEMEQEFPEEFARTQESVFRSGTDISVTNSLYHYYALITGRAVQQEKAKVL
                YVDTTSYTGLNLLPELRKRRNYDFFCLNDGSFPEVPATERAERVVSFLERYFPIPAPW
                EKVATDFNRQDFASPTVSAPLEDGQTANPAQTAR"
CDS             complement(8509. .8724)
                /codon_start=1
                /transl_table=11
                /product="u2266d"
                /protein_id="AAA62978.1"
                /db_xref="GI:699214"
                /translation="MTIYPNAAPDVWVAISQAMTTAHTINHQQRYLDDVLLAATGFGN
                MGGIDIRTERVIPGAWYPRSGHLQPAA"
CDS             complement(8721. .8795)
                /codon_start=1
                /transl_table=11
                /product="u2266e"
                /protein_id="AAA62977.1"
                /db_xref="GI:699213"
                /translation="MLLDDVNLQLPQNQAKKTTAGGTG"
CDS             9802. .11472
                /codon_start=1
                /transl_table=11
                /product="purF"
                /protein_id="AAA62988.1"
                /db_xref="GI:699224"
                /translation="MCLAVGVGVRAPKHVPQIRRLGRAGRRLRCVTNCALGSCPIVTV
                QQPGRDFSSPREECGVFGVWAPGELVAKLTYFGLYALQHRGQEAAGIAVADGSQVLVF
                KDLGLVSQVFDEQTLAAMEGHVAIGHCRYSTTGDTTWENAQPVFRNIAAGSGVALGHN
                GNLVNTAELAARARDAGLIAKRCPAPATTDSDILGALLAHGAADSTLEQAALELLPTV
                RGAFCLTFMDENTLYACRDPYGVRPLSLGRLDRGWVVASETAGLDIVGASFVRDIEPG
                ELLAIDADGVRSTRFANPTPKGCVFEYVYLARPDSTLAGRSVHGTRVEIGRRLARECP
                VEADLVIGVPESGTPAAVGYAQESGISYGQGLMKNAYVGRTFIQPSQTIRQLGIRLKL
                NPLKEVIRGKRLIVVDDSVVRGNTQRALVRMLREAGAVELHVRIASPPVKWPCFYGID
                FPSPAELIANVVADEEEMLEAVRQGIGADTLGYISLRGMIAASEQPASRLCYACFDGR
                YPIELPSEAMLGKNVIEHMLANAARGAGLRDLAADQVPVDADENCVWR"
CDS             11547. .12140
                /codon_start=1
                /transl_table=11
                /product="purM"
                /protein_id="AAA62987.1"
                /db_xref="GI:699223"
                /translation="MTTGSLYRDDGSRKKPRSQLGNHGITYASTGVDIEAGDRSVELF
                KPMAVKAIRTEVLGGLGGFAGLFALRSGYREPVLAASTDGVGTKLAVAQAMDKHATVG
                LDLVAMVVDDLVVCGAEPLFLQDYIAVGRTVPERLSAIVDGIAEGCMLAGCALLGGET
                AEHPGLMEVDHYDIAATGVGVVEADDTWGQIVSNPVT"
CDS             12218. .12664
                /codon_start=1
                /transl_table=11
                /product="purX"
                /protein_id="AAA62985.1"
                /db_xref="GI:699221"
                /translation="MNLAGYVEEFGRTLGEELLEPTRIYAKDCLALAAETHVRTFCHV
                TGGGLAGNLQRVIPHGLVAEVDRGTWTPAPVFAMIAQRGRVAREEMEKTFNMGVGMIA
                VVAPEDTDRALAILTARHLDCWVLGTVCKGSKVGPRAKLVGQHPRF"
CDS             complement(13113. .14213)
                /codon_start=1
                /transl_table=11
                /product="u2266f"
                /protein_id="AAA62976.1"
                /db_xref="GI:699212"
                /translation="MNAVLAPDPGPDAGAIWHFGDPLGEQRSAETDAVLVDRSHRATL
                TLIGSDRQTWLHSISTQHVSALPDGASTQNLSLDSQGRVEDHWIQTELGGTTYLDTEP
                WRGEPLLEYLRKMVFWSEVTPGSADMAMLSLLGPRLTEQAVHNALYLDALPDESTAVP
                LTAGSFVRRIPGAPAGQIELDLLVPRRESADWQSRLAGVGVRSAGMWAYEAHRVAALR
                PRLGVDTDQRTIPHEVGWIGGPGEGAVHLDKGCYRGQETVARVQNLGKPPRMLVLLHL
                DGSVERTSTGDAVLANSGAVGRLGTVVEHVDLGPVALALLKRGLPTDTELTIELEGPQ
                DGSVAAVIDPDSLPAADEIGVGRLAIERLRGR"
CDS             15719. .15751
                /codon_start=1
                /transl_table=11
                /product="u2266g"
                /protein_id="AAA62975.1"
                /db_xref="GI:699211"
                /translation="MDLSNCEGRL"
CDS             complement(15757. .15912)
                /codon_start=1
                /transl_table=11
                /product="u2266h"
                /protein_id="AAA62974.1"
                /db_xref="GI:699210"
                /translation="MSMLAATTRQYAASSCTLRDNTLRVYEAAACSGQTPMPATEQGK
                GLSLDVQ"
CDS             complement(17941. .18669)
                /codon_start=1
                /transl_table=11
                /product="u2266i"
                /protein_id="AAA62973.1"
                /db_xref="GI:699209"
                /translation="MGGTSTRLRGRAVHVTSDEVRDGAGSPADSSKGNKCTAAGMFQA

 Query Match          42.1%;  Score 376.8;  DB 1;  Length 40123;
 Best Local Similarity  67.4%;  Pred. No. 1.3e-76;
 Matches  531;  Conservative    0;  Mismatches  257;  Indels    0;  Gaps    0;

Qy      96 AAGCGATGTCGAAGCTCAAGCTCAATGATGTCAACATCTACTACGGTGATTTCCACGCAG 155
           ||| |    ||   |  | ||||| |  |||||||| || |||||   ||| || ||||
Db   27034 AAGTGGCCAAGAGATTGGATCTCAAAGGCGTCAACATTTATTACGGATCTTTTCAAGCAG 27093

Qy     156 TGCAGAACGTGAACCTCGAGGTTCCTGCACGCTCTGTCACCGCATTCATCGGACCATCCG 215
           |   | | |||   || | ||| |   |  |    || || || |||||||| | || |
Db   27094 TTTTGGATGTGTCGCTGGCGGTACTGCCCAGAAGCGTGACGGCGTTCATCGGAGCCTCAG 27153

Qy     216 GCTGTGGCAAGTCCACAGTTCTCCGCTCCATCAATCGTATGCACGAGGTCACCCCAGGTG 275
           |||| |||||| | || ||  | ||| || | ||||| ||||| ||||||  ||| || |
Db   27154 GCTGCGGCAAGACGACGGTGTTGCGCACCTTAAATCGGATGCATGAGGTCGTCCCTGGCG 27213

Qy     276 CATACGTCAAGGGCGAGATCCTTCTCGACGGCGAGAACATCTACGGCTCCAAGATCGACC 335
           |   ||  |||||    |  | |||||||  ||  |||| || | | |    ||||| |
Db   27214 CCCGTGTTGAGGGCACCGTGTTACTCGACGATGAAGACATTTATGCCCCGGGTATCGATC 27273

Qy     336 CAGTTGCAGTCCGTAACACCATCGGCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCA 395
           | || |  || ||     | ||||| ||||| ||||||  | | ||||| |||||  |||
Db   27274 CCGTCGGTGTGCGACGGGCGATCGGGATGGTGTTCCAGCGGCCCAACCCGTTCCCCGCCA 27333

Qy     396 TGTCCATCGAGGACAACGTGGTTGCAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGA 455
           |||| ||    || |||||||| || ||| |||||||    |||| |   ||   |||
Db   27334 TGTCAATTCGCGATAACGTGGTGGCCGGTTTGAAGCTGCAAGGCGTGCGTAATCGTAAGG 27393

Qy     456 AGCTCAAGGAAGTTGCTGAGAAGTCTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGG 515
           |||  | ||    || ||  | | | | || ||  | || |||||||| ||||| ||||
Db   27394 TGCTTGATGATACCGCAGAATACTTTTTGCGCGGTACGAATCTGTGGGATGAGGTCAAGG 27453

Qy     516 ATCGTCTGGACAAGCCAGGCGGCGGCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCG 575
           | ||  |||| ||||| ||||| || || ||||| || |||||||| || ||||| ||||
Db   27454 ACCGATTGGATAAGCCTGGCGGTGGACTGTCCGGGGGGCAGCAGCAACGGCTGTGTATCG 27513

Qy     576 CTCGCGCGATCGCGGTTGAGCCAGAAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTG 635
           |||| || || || ||  | |||||  |  | |||||||||||||| |||||| |||| |
Db   27514 CTCGGGCCATTGCCGTGCAACCAGACGTGTTACTCATGGACGAGCCGTGCTCCTCGCTCG 27573

Qy     636 ACCCAATCTCAACCCTGGCTGTGGAGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCA 695
           | || || || ||  ||||| | || || || |||   ||  ||||  | |||| |||||
Db   27574 ATCCCATTTCGACGATGGCTATCGAAGAACTGATCGGTGAATTGAAACAGGAGTACACCA 27633

Qy     696 TCGTCATCGTGACCCACAACATGCAGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCT 755
           | || || ||||||||||||||||||| || || ||| |||||||| |||||||
Db   27634 TTGTAATAGTTACCCACAACATGCAGCAGGCGGCGCGGGTGAGCGATCAGACTGCGTTCT 27693

Qy     756 ACTCCCTGGAGGCGACCGGTAGGCCAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGA 815
           |  | |||| ||    || || || || |||| || || |     ||| |||| |
Db   27694 TCAACTTGGAAGCCGTAGGAAGACCCGGCCGGCTGGTGGAGATAGATGACACCGAGAAAA 27753
```

```
Qy    816 TCTTCGAAAACCCAGATCAGAAGGAAACCGAAGATTACATCTCCGGCCGCTTCGGATAAA 875
          ||||    || ||     | ||||  || || || ||||||||||||||||||| ||
Db  27754 TCTTTTCCAATCCGACCGAAAAGGCCACTGAGGACTACATCTCCGGCCGCTTCGGCTAGT 27813

Qy    876 TCGAAGAA 883
          ||  ||||
Db  27814 TCATAGAA 27821


RESULT 11
MLEPRTN8/c
LOCUS       MLEPRTN8              342300 bp    DNA     linear   BCT 22-FEB-2001
DEFINITION  Mycobacterium leprae strain TN complete genome; segment 8/10.
ACCESSION   AL583924 AL450380
VERSION     AL583924.1  GI:13093618
KEYWORDS    .
SOURCE      Mycobacterium leprae.
  ORGANISM  Mycobacterium leprae
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium.
REFERENCE   1  (bases 1 to 342300)
  AUTHORS   Cole,S.T., Eiglmeier,K., Parkhill,J., James,K.D., Thomson,N.R.,
            Wheeler,P.R., Honore,N., Ganier,T., Churcher,C., Harris,D.,
            Mungall,K., Basham,D., Brown,D., Chillingworth,T., Connor,R.,
            Davies,R.M., Devlin,K., Duthoy,S., Feltwell,T., Fraser,A.,
            Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Lacroix,C.,
            Maclean,J., Moule,S., Murphy,L., Oliver, Quail,M.A.,
            Rajandream,M.-A., Rutherford,K.M., Rutter,S., Seeger,K., Simon,S.,
            Simmonds,M., Skelton,J., Squares,R., Squares,S., Stevens,K.,
            Taylor,K., Whitehead,S., Woodward,J.R. and Barrell,B.G.
  TITLE     Massive gene decay in the leprosy bacillus
  JOURNAL   Nature 409 (6823), 1007-1011 (2001)
  MEDLINE   21128732
   PUBMED   11234002
REFERENCE   2  (bases 1 to 342300)
  AUTHORS   Parkhill,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2001) Submitted on behalf of the Mycobacterium
            leprae sequencing teams, The Sanger Centre, Wellcome Trust Genome
            Campus, Hinxton, Cambridge, CB10 1SA, UK Unitie de Genetique
            Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Rouux,
            75724, Paris Cedex, France. E-mail: parkhill@sanger.ac.uk
COMMENT     Notes:
            Details of M. leprae sequencing at the Sanger Centre are available
            from http://www.sanger.ac.uk/Projects/M_leprae/ A relational
            datbase containing the M. leprae sequences is available from
            http://genolist.pasteur.fr/Leproma/.
FEATURES             Location/Qualifiers
     source          1. .342300
                     /organism="Mycobacterium leprae"
                     /strain="TN"
                     /db_xref="taxon:1769"
     gene            275. .430
                     /gene="ML1963"
     CDS             275. .430
                     /gene="ML1963"
                     /note="Possible pseudogene of M. tuberculosis orthologue
                     Rv3463 (Best blastx score 121)"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="probable neuraminidase (pseudogene)"
     gene            483. .1484
                     /gene="rmlB"
                     /note="ML1964"
     CDS             483. .1484
                     /gene="rmlB"
                     /note="Similar to M. tuberculosis dTDP-glucose
                     4,6-dehydratase rmlB Rv3464 TR:O06329 (EMBL:Z95390) (331
                     aa); Fasta score E(): 0, 84.0% identity in 331 aa overlap,
                     and TR:Q50556 (EMBL:U43540) (329 aa); Fasta score E(): 0,
                     78.2% identity in 331 aa overlap, and to others e.g.
                     Streptococcus pneumoniae DTDP-glucose-4,6-dehydratase cpsN
                     TR:O54611 (EMBL:AF030364) (349 aa); Fasta score E(): 0,
                     59.9% identity in 334 aa overlap. Contains Pfam match to
                     entry PF01370 Epimerase, NAD dependent
                     epimerase/dehydratase family.
                     Similar to ML0204, ML0751, ML1942 and ML2428"
                     /codon_start=1
                     /transl_table=11
                     /product="dTDP-glucose 4,6-dehydratase"
                     /protein_id="CAC30919.1"
                     /db_xref="GI:13093619"
                     /translation="MTMRLLVTGGAGFIGANFVHNAVRDHPDDTVTVLDAMTYAGRHE
                     SLADVDHAVRLVQGNITDTKLVFRLVAESDVVVHFAAETHVDNALDDPEPFLRTNVIG
                     TFTILEAVRHHSVRLHHISTDEVYGDLELDEPTRFTESTPYNPSSPYSATKAAGDMLV
                     RAWVRSYGVRATISNCSNNYGPYQHVEKFIPRQITNVLTGRRPKLYGTGVNVRDWIHV
                     DDHNSAVWQILEKGQIGRTYLIGAENGHNNLTVLRTLLQMMGRDPNDFDHVTDRVGHD
                     LRYAIDPTPLYNELCWAPKHANFDEGLRATIDWYRNNESWWRPLKDAVEARYEERGR"
     misc_feature    495. .1424
                     /gene="rmlB"
                     /note="Pfam match to entry PF01370 Epimerase, NAD
                     dependent epimerase/dehydratase family, score 470.30,
                     E-value 1.6e-137"
     gene            1486. .2094
                     /gene="rmlC"
                     /note="ML1965"
     CDS             1486. .2094
                     /gene="rmlC"
                     /note="Similar to M. tuberculosis dTDP-4-dehydrorhamnose
                     3,5-epimerase rmlC Rv3465 TR:O06330 (EMBL:Z95390) (202
                     aa); Fasta score E(): 0, 75.4% identity in 199 aa overlap,
                     and to many others e.g. Streptomyces griseus
                     DTDP-4-dehydrorhamnose 3,5-epimerase SW:STRM_STRGR
                     (P29783) (200 aa); Fasta score E(): 0, 45.6% identity in
                     193 aa overlap. Contains Pfam match to entry PF00908
                     dTDP_sugar_isom, dTDP-4-dehydrorhamnose 3,5-epimerase."
                     /codon_start=1
                     /transl_table=11
                     /product="dTDP-4-dehydrorhamnose 3,5-epimerase"
                     /protein_id="CAC30920.1"
                     /db_xref="GI:13093620"
                     /translation="MNVKELDISGAWEITPTIHSDSRGMFFEWLTASRFSAFTGHRLD
                     VRQANCSVSSAGVLRGLHFAQVPPSQAKYVTCVTGSVFDVVVDIRLGSPTFGCWTSIT
                     LNDSNRKTIYISESLAHGFLALQDNSTVMYLCSAEYNPAREHAICATDPDLAIDWPVV
                     DGVDTHLSKRDATAPNLDDIRASGLLPTWDETQNFIANLGSK"
     misc_feature    1492. .2025
                     /gene="rmlC"
                     /note="Pfam match to entry PF00908 dTDP_sugar_isom,
                     dTDP-4-dehydrorhamnose 3,5-epimerase, score 308.80,
                     E-value 6.5e-89"
     gene            2139. .2624
                     /gene="lpqH"
                     /note="ML1966"
     CDS             2139. .2624
                     /gene="lpqH"
                     /note="Similar to M. tuberculosis 19 kDa lipoprotein
                     antigen precursor lpqH Rv3763. Contains a possible
                     N-terminal signal sequence and a PS00013 Prokaryotic
                     membrane lipoprotein lipid attachment site."
                     /codon_start=1
                     /transl_table=11
                     /product="possible lipoprotein"
                     /protein_id="CAC30921.1"
                     /db_xref="GI:13093621"
                     /translation="MRHKLLAAIYAVTIMAGAAGCSGGTQAPTPSVSKTTNSSPTTVA
                     SSIPDAAAGETKVTIGGQPQKVSGPVVCSTTNGKFSIAIGDMITGVIVGLEPDASVVH
                     NAGLGTIDGVVIAFTEGVPSENANATKNGNTYQITGTASGVDNTGQQIHKSFEIEVTC
                     R"
     misc_feature    2169. .2201
                     /gene="lpqH"
                     /note="PS00013 Prokaryotic membrane lipoprotein lipid
                     attachment site"
     gene            complement(4768. .4995)
```

```
                     /gene="ML1967"
     CDS             complement(4768. .4995)
                     /gene="ML1967"
                     /note="Possible pseudogene of M. tuberculosis paralog PPE"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="PPE family protein (pseudogene)"
     gene            complement(5464. .5649)
                     /gene="ML1968"
     CDS             complement(5464. .5649)
                     /gene="ML1968"
                     /note="Possible pseudogene of M. tuberculosis paralog PE"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="PE family protein (pseudogene)"
     repeat_region   complement(5661. .8043)
                     /note="Dispersed repeat, LEPREP, copy 6"
     gene            complement(5848. .5946)
                     /gene="ML1969"
     CDS             complement(5848. .5946)
                     /gene="ML1969"
                     /note="Similar to Agrobacterium tumefaciens transposase
                     TR:Q44454 (EMBL:Z18270) (366 aa) fasta scores: E():
                     0.0002, 51.5% id in 33 aa, and to Pseudomonas putida
                     transposase tnpa1 TR:Q9R9U9 (EMBL:AJ245436) (355 aa) fasta
                     scores: E(): 0.00088, 48.5% id in 33 aa"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
     gene            complement(6211. .7354)
                     /gene="ML1970"
     CDS             complement(6211. .7354)
                     /gene="ML1970"
                     /note="Similar to many e.g. Cryphonectria parasitica
                     (Chesnut blight fungus) putative maturase TR:AAF27656
                     (EMBL:AF218567) (778 aa) fasta scores: E(): 7.8e-11, 29.2%
                     id in 216 aa"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="putative group II intron maturase-related
                     protein"
     stem_loop       complement(6674. .6727)
                     /gene="ML1970"
     gene            complement(7559. .7698)
                     /gene="ML1971"
     CDS             complement(7559. .7698)
                     /gene="ML1971"
                     /note="Similar to Pseudomonas putida transposase tnpa1
                     TR:Q9R9U9 (EMBL:AJ245436) (355 aa) fasta scores: E():
                     3.4e-05, 47.8% id in 46 aa, and to Agrobacterium
                     tumefaciens transposase TR:Q44454 (EMBL:Z18270) (366 aa)
                     fasta scores: E(): 0.00026, 41.3% id in 46 aa"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
     repeat_region   complement(8044. .8189)
                     /note="146 bp sequence found at 5'-end of copies 6 and 7
                     of LEPREP"
     gene            complement(8461. .8691)
                     /gene="ML1972"
     CDS             complement(8461. .8691)
                     /gene="ML1972"
                     /note="unknown."

  Query Match             42.1%;  Score 376.8;  DB 1;  Length 342300;
  Best Local Similarity   67.4%;  Pred. No. 1.7e-76;
  Matches  531;  Conservative    0;  Mismatches  257;  Indels    0;  Gaps    0;

Qy      96 AAGCGATGTCGAAGCTCAAGCTCAATGATGTCAACATCTACTACGGTGATTTCCACGCAG 155
           ||| |     ||   | | ||||| |  |||||||| || |||||   ||| || ||||

Db 257201 AAGTGGCCAAGAGATTGGATCTCAAAGGCGTCAACATTTATTACGGATCTTTTCAAGCAG 257142

Qy     156 TGCAGAACGTGAACCTCGAGGTTCCTGCACGCTCTGTCACCGCATTCATCGGACCATCCG 215
           |   | | |||   || | ||| |   |  |    || || || |||||||||| | || |
Db 257141 TTTTGGATGTGTCGCTGGCGGTACTGCCCAGAAGCGTGACGGCGTTCATCGGAGCCTCAG 257082

Qy     216 GCTGTGGCAAGTCCACAGTTCTCCGCTCCATCAATCGTATGCACGAGGTCACCCCAGGTG 275
           |||| |||||| | || ||  | ||| || | ||||| ||||| ||||||  ||| || |
Db 257081 GCTGCGGCAAGACGACGGTGTTGCGCACCTTAAATCGGATGCATGAGGTCGTCCCTGGCG 257022

Qy     276 CATACGTCAAGGGCGAGATCCTTCTCGACGGCGAGAACATCTACGGCTCCAAGATCGACC 335
           |    ||  |||||    |  | |||||||  ||  |||| || | | |    ||||| |
Db 257021 CCCGTGTTGAGGGCACCGTGTTACTCGACGATGAAGACATTTATGCCCCGGGTATCGATC 256962

Qy     336 CAGTTGCAGTCCGTAACACCATCGGCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCA 395
           | || |  || ||     | ||||| ||||| ||||||  | | |||| |||||  |||
Db 256961 CCGTCGGTGTGCGACGGGCGATCGGGATGGTGTTCCAGCGGCCCAACCCGTTCCCCGCCA 256902

Qy     396 TGTCCATCGAGGACAACGTGGTTGCAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGA 455
           |||| ||    || |||||||| || ||| |||||||    |||| |   ||    |||
Db 256901 TGTCAATTCGCGATAACGTGGTGGCCGGTTTGAAGCTGCAAGGCGTGCGTAATCGTAAGG 256842

Qy     456 AGCTCAAGGAAGTTGCTGAGAAGTCTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGG 515
           |||  | ||    || ||  | | | || ||  | || |||||||| ||||| ||||
Db 256841 TGCTTGATGATACCGCAGAATACTTTTTGCGCGGTACGAATCTGTGGGATGAGGTCAAGG 256782

Qy     516 ATCGTCTGGACAAGCCAGGCGGCGGCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCG 575
           | ||  |||| ||||| ||||| || || ||||| || |||||||| || ||||| ||||
Db 256781 ACCGATTGGATAAGCCTGGCGGTGGACTGTCCGGGGGGCAGCAGCAACGGCTGTGTATCG 256722

Qy     576 CTCGCGCGATCGCGGTTGAGCCAGAAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTG 635
           |||| || || || ||  | |||||   |  | ||||||||||||||| |||||| |||| |
Db 256721 CTCGGGCCATTGCCGTGCAACCAGACGTGTTACTCATGGACGAGCCGTGCTCCTCGCTCG 256662

Qy     636 ACCCAATCTCAACCCTGGCTGTGGAGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCA 695
           | || || || ||  ||||| | || || || |||   ||  ||||  | |||| |||||
Db 256661 ATCCCATTTCGACGATGGCTATCGAAGAACTGATCGGTGAATTGAAACAGGAGTACACCA 256602

Qy     696 TCGTCATCGTGACCCACAACATGCAGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCT 755
           | || || || ||||||||||||||||||| || || |||  |||||||| |||||||
Db 256601 TTGTAATAGTTACCCACAACATGCAGCAGGCGGCGCGGGTGAGCGATCAGACTGCGTTCT 256542

Qy     756 ACTCCCTGGAGGCGACCGGTAGGCCAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGA 815
           |  | |||| ||    || || || || || ||||| || || |     ||| |||| |
Db 256541 TCAACTTGGAAGCCGTAGGAAGACCCGGCCGGCTGGTGGAGATAGATGACACCGAGAAAA 256482

Qy     816 TCTTCGAAAACCCAGATCAGAAGGAAACCGAAGATTACATCTCCGGCCGCTTCGGATAAA 875
           ||||    || ||     | ||||  || || || ||||||||||||||||||| ||
Db 256481 TCTTTTCCAATCCGACCGAAAAGGCCACTGAGGACTACATCTCCGGCCGCTTCGGCTAGT 256422

Qy     876 TCGAAGAA 883
           ||  ||||
Db 256421 TCATAGAA 256414


RESULT 12
MSGL611CS
LOCUS       MSGL611CS             37769 bp    DNA     linear   BCT 17-DEC-2001
DEFINITION  Mycobacterium leprae cosmid L611 DNA sequence.
ACCESSION   L78822
VERSION     L78822.1  GI:1377787
KEYWORDS    .
SOURCE      Mycobacterium leprae (clone: cosmid L611) (tissue library: Lorist
            6) DNA.
  ORGANISM  Mycobacterium leprae
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium.
REFERENCE   1  (sites)
  AUTHORS   Eiglmeier,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T.
  TITLE     Use of an ordered cosmid library to deduce the genomic organization
            of Mycobacterium leprae
  JOURNAL   Mol. Microbiol. 7 (2), 197-206 (1993)
```

```
  MEDLINE   93188700
   PUBMED   8446027
REFERENCE   2  (bases 1 to 37769)
  AUTHORS   Smith,D.R., Richterich,P., Rubenfield,M., Rice,P.W., Butler,C.,
            Lee,H.M., Kirst,S., Gundersen,K., Abendschan,K., Xu,Q., Chung,M.,
            Deloughery,C., Aldredge,T., Maher,J., Lundstrom,R., Tulig,C.,
            Falls,K., Imrich,J., Torrey,D., Engelstein,M., Breton,G., Madan,D.,
            Nietupski,R., Seitz,B., Mao,J.I. et al.
  TITLE     Multiplex sequencing of 1.5 Mb of the Mycobacterium leprae genome
  JOURNAL   Genome Res. 7 (8), 802-819 (1997)
  MEDLINE   97413161
   PUBMED   9267804
COMMENT     This sequence was generated by the Genome Sequencing Center at
            Genome Therapeutics Corporation (Collaborative Research Division),
            100
            Beaver St., Waltham, MA, 02154.  Please contact Doug Smith
            (smith@cric,com) for further information.  The sequence represents
            the
            insert of a Lorist 6 cosmid clone from a mapped set of clones
            constructed from M. leprae genomic DNA isolated from armadillo
            liver
            [3].  The sequence may not represent the entire cloned insert of
            the
            cosmid if an overlapping region was previously sequenced from
            another
            clone.  Coding sequences larger than 60 amino acids were predicted
            on
            the basis of codon usage and homology information.  An attempt was
            made
            to locate the most probable start site based on codon usage,
            homology,
            the presence of a Shine-Dalgarno sequence, or overlapping orf that
            suggested translational coupling.  It is possible that the actual
            start
            site differs from the one selected.
FEATURES             Location/Qualifiers
     source          1. .37769
                     /organism="Mycobacterium leprae"
                     /specific_host="Dasypus novemcinctus"
                     /db_xref="taxon:1769"
                     /clone="cosmid L611"
                     /tissue_lib="Lorist 6"
                     /note="The liver of the armadillo was used to isolate the
                     Mycobacterium leprae."
BASE COUNT     7786 a   9706 c  11113 g   9164 t
ORIGIN

 Query Match           42.0%;  Score 375.2;  DB 1;  Length 37769;
 Best Local Similarity  67.3%;  Pred. No. 3.1e-76;
  Matches  530;  Conservative    0;  Mismatches  258;  Indels    0;  Gaps    0;

Qy      96 AAGCGATGTCGAAGCTCAAGCTCAATGATGTCAACATCTACTACGGTGATTTCCACGCAG 155
           ||| |     ||   |  | ||||| |  |||||||| || |||||   ||| || ||||
Db    9337 AAGTGGCCAAGAGATTGGATCTCAAAGGCGTCAACATTTATTACGGATCTTTTCAAGCAG 9396

Qy     156 TGCAGAACGTGAACCTCGAGGTTCCTGCACGCTCTGTCACCGCATTCATCGGACCATCCG 215
           |   | | |||   || | ||| |   |  |    || || || ||||||||| | || |
Db    9397 TTTTGGATGTGTCGCTGGCGGTACTGCCCAGAAGCGTGACGGCGTTCATCGGAGCCTCAG 9456

Qy     216 GCTGTGGCAAGTCCACAGTTCTCCGCTCCATCAATCGTATGCACGAGGTCACCCCAGGTG 275
           |||| |||||| | || ||  | ||| || | ||||| ||||| ||||||  ||| || |
Db    9457 GCTGCGGCAAGACGACGGTGTTGCGCACCTTAAATCGGATGCATGAGGTCGTCCCTGGCG 9516

Qy     276 CATACGTCAAGGGCGAGATCCTTCTCGACGGCGAGAACATCTACGGCTCCAAGATCGACC 335
           |    ||  |||||    |  | |||||||  ||  |||| || | | |    |||||
Db    9517 CCCGTGTTGAGGGCACCGTGTTACTCGACGATGAAGACATTTATGCCCCGGGTATCGATT 9576

Qy     336 CAGTTGCAGTCCGTAACACCATCGGCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCA 395
           | || |  || ||     | ||||| ||||| ||||||  | | ||||| |||||  |||
Db    9577 CCGTCGGTGTGCGACGGGCGATCGGGATGGTGTTCCAGCGGCCCAACCCGTTCCCCGCCA 9636

Qy     396 TGTCCATCGAGGACAACGTGGTTGCAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGA 455
           |||| ||    || |||||||| || ||| |||||||    |||| |   ||    |||

Db    9637 TGTCAATTCGCGATAACGTGGTGGCCGGTTTGAAGCTGCAAGGCGTGCGTAATCGTAAGG 9696

Qy     456 AGCTCAAGGAAGTTGCTGAGAAGTCTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGG 515
           |||  | ||    || ||  | | | | || ||  | || |||||||| ||||| ||||
Db    9697 TGCTTGATGATACCGCAGAATACTTTTTGCGCGGTACGAATCTGTGGGATGAGGTCAAGG 9756

Qy     516 ATCGTCTGGACAAGCCAGGCGGCGGCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCG 575
           | ||  |||| ||||| ||||| || || ||||| || |||||||| || ||||| ||||
Db    9757 ACCGATTGGATAAGCCTGGCGGTGGACTGTCCGGGGGGCAGCAGCAACGGCTGTGTATCG 9816

Qy     576 CTCGCGCGATCGCGGTTGAGCCAGAAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTG 635
           |||| || || || ||  | |||||  |  | ||||||||||||||| |||||| |||| |
Db    9817 CTCGGGCCATTGCCGTGCAACCAGACGTGTTACTCATGGACGAGCCGTGCTCCTCGCTCG 9876

Qy     636 ACCCAATCTCAACCCTGGCTGTGGAGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCA 695
           | || || || ||  ||||| | || || || |||   ||  ||||  | |||| |||||
Db    9877 ATCCCATTTCGACGATGGCTATCGAAGAACTGATCGGTGAATTGAAACAGGAGTACACCA 9936

Qy     696 TCGTCATCGTGACCCACAACATGCAGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCT 755
           | || || || |||||||||||||||||| || || ||| |||||||| |||||||
Db    9937 TTGTAATAGTTACCCACAACATGCAGCAGGCGGCGCGGGTGAGCGATCAGACTGCGTTCT 9996

Qy     756 ACTCCCTGGAGGCGACCGGTAGGCCAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGA 815
           |  | |||| ||    || || || || || ||||| || || |     ||| |||| |
Db    9997 TCAACTTGGAAGCCGTAGGAAGACCCGGCCGGCTGGTGGAGATAGATGACACCGAGAAAA 10056

Qy     816 TCTTCGAAAACCCAGATCAGAAGGAAACCGAAGATTACATCTCCGGCCGCTTCGGATAAA 875
           ||||    || ||     | ||||  || || || ||||||||||||||||||| ||
Db   10057 TCTTTTCCAATCCGACCGAAAAGGCCACTGAGGACTACATCTCCGGCCGCTTCGGCTAGT 10116

Qy     876 TCGAAGAA 883
           ||  ||||
Db   10117 TCATAGAA 10124


RESULT 13
AE001862
LOCUS       AE001862               213732 bp    DNA     linear   BCT 22-NOV-1999
DEFINITION  Deinococcus radiodurans R1 section 1 of 2 of the complete
            chromosome 2.
ACCESSION   AE001862 AE001825
VERSION     AE001862.1  GI:6460468
KEYWORDS    .
SOURCE      Deinococcus radiodurans.
  ORGANISM  Deinococcus radiodurans
            Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
            Deinococcaceae; Deinococcus.
REFERENCE   1  (bases 1 to 213732)
  AUTHORS   White,O., Eisen,J.A., Heidelberg,J.F., Hickey,E.K., Peterson,J.D.,
            Dodson,R.J., Haft,D.H., Gwinn,M.L., Nelson,W.C., Richardson,D.L.,
            Moffat,K.S., Qin,H., Jiang,L., Pamphile,W., Crosby,M., Shen,M.,
            Vamathevan,J.J., Lam,P., McDonald,L., Utterback,T., Zalewski,C.,
            Makarova,K.S., Aravind,L., Daly,M.J., Fraser,C.M. et al.
  TITLE     Genome sequence of the radioresistant bacterium Deinococcus
            radiodurans R1
  JOURNAL   Science 286 (5444), 1571-1577 (1999)
  MEDLINE   20036896
   PUBMED   10567266
REFERENCE   2  (bases 1 to 213732)
  AUTHORS   White,O., Eisen,J.A., Heidelberg,J.F., Hickey,E.K., Peterson,J.D.,
            Dodson,R.J., Haft,D.H., Gwinn,M.L., Nelson,W.C., Richardson,D.L.,
            Moffat,K.S., Qin,H., Jiang,L., Pamphile,W., Crosby,M., Shen,M.,
            Vamathevan,J.J., Lam,P., McDonald,L., Utterback,T., Zalewski,C.,
            Makarova,K.S., Aravind,L., Daly,M.J., Minton,K.W.,
            Fleischmann,R.D., Ketchum,K.A., Nelson,K.E., Salzberg,S.,
            Smith,H.O., Venter,J.C. and Fraser,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1999) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .213732
                     /organism="Deinococcus radiodurans"
                     /strain="R1"
```

```
                /db_xref="taxon:1299"
                /chromosome="2"
     gene       653. .1435
                /gene="DRA0001"
     CDS        653. .1435
                /gene="DRA0001"
                /note="similar to GB:D26185 SP:P37522 PID:467381
                PID:580906 GB:AL009126 percent identity: 54.96; identified
                by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="chromosome partitioning ATPase, putative, ParA
                family"
                /protein_id="AAF12301.1"
                /db_xref="GI:6460595"
                /translation="MVSAVKTLTVFNHAGGAGKTSLTLNVGYELARGGLRVLLLDLDP
                QANLTGWLGISGVTREMTVYPVAVDGQPLPSPVKAFGLDVIPAHVSLAVAEGQMMGRV
                GAQGRLRRALAEVSGDYDVALIDSPPSLGQLAILAALAADQMIVPVPTRQKGLDALPG
                LQGALTEYREVRPDLTVALYVPTFYDARRRHDQEVLADLKAHLSPLARPVPQREAVWL
                DSTAQGAPVSEYAPGTPVHADVQRLTADIAAAIGVAYPGENA"
     gene       1432. .2313
                /gene="DRA0002"
     CDS        1432. .2313
                /gene="DRA0002"
                /note="similar to GB:L34077 GB:X85964 PID:407375
                PID:757761 PID:1163135 percent identity: 57.34; identified
                by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="chromosome partitioning protein, ParB family"
                /protein_id="AAF12300.1"
                /db_xref="GI:6460594"
                /translation="MTRRRPERRRDLLGLLGETPVDLSQANDIRALPVNELKVGSTQP
                RRSFDLERLSELAESIRAHGVLQPLLVRSVDGQYEIVAGERRWRAAQLAGLAEVPVVV
                RQLSNEQARAAALIENLQRDNLNVIDEVDGKLELIALTLGLEREEARKRLMQLLRAVP
                GDEHEQLDQVFRSMGETWRTFAKNKLRILNWPQPVLEALRAGLPLTLGSVVASAPPER
                QAELLKLAQNGASRSQLLQALQTPSQTSAVTPEHFAKVLSSKRFLSGLDTPTREALDR
                WLARMPERVRQAIDEQS"
     gene       complement(2418. .3881)
                /gene="DRA0003"
     CDS        complement(2418. .3881)
                /gene="DRA0003"
                /note="similar to GB:M88334 SP:P25526 PID:147901 GB:U00096
                PID:1789015 percent identity: 69.60; identified by
                sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="succinate-semialdehyde dehydrogenase"
                /protein_id="AAF12294.1"
                /db_xref="GI:6460588"
                /translation="MTTTDLRTTYSSVTRSQAYFDGEWRNAPRNFEVRHPGNGEVIGE
                VADCTPTDARQAIDAAEVALREWRQVNPYERGKILRRWHDLMFEHKEELAQLMTLEMG
                KPISETRGEVHYAASFIEWCAEEAGRIAGERINLRFPHKRGLTISEPVGIVYAVTPWN
                FPAGMITRKAAPALAAGCVMILKPAELSPMTALYLTELWLKAGGPANTFQVLPTNDAS
                ALTQPFMNDSRVRKLTFTGSTEVGRLLYQQAAGTIKRVSLELGGHAPFLVFDDADLER
                AASEVVASKFRNSGQTCVCTNRVYVQRGVAEEFIRLLTEKTAALQLGDPFDEATQVGP
                VVEQAGLDKVQRQVQDALTKGAQATTGGQVSSGLFFQPTVLVDVAPDSLILREETFGP
                VAPVTIFDTEEEGLRLANDSEYGLAAYAYTRDLGRAFRIAEGLEYGIVGINDGLPSSA
                APHVPFGGMKNSGVGREGGHWGLEEYLETKFVSLGLS"
     gene       4024. .4668
                /gene="DRA0004"
     CDS        4024. .4668
                /gene="DRA0004"
                /note="similar to GP:2924378 percent identity: 55.29;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAF12299.1"
                /db_xref="GI:6460593"
                /translation="MNSLSSLAAVLLMAGRSTRMGQPKALLPLGGMPLCRHAAQTLVE
                AGYGEVWAVLPPGETGDAVRAALQGLPLRWVTNPQPAQGLLSSFQAAGEALWAESAAP
                AVVFTLADMPLVRRDTHQALGTVFGAEQPAAIITRYGEVAAPPTLLRRDLLPRLHEFP
                PSDHGPRTLLRELGRELKKIDRPAAELLDIDTPEAWAQAQQLFAEVTPGAQDGP"
     gene       complement(4719. .5897)
                /gene="DRA0005"
     CDS        complement(4719. .5897)
                /gene="DRA0005"
                /note="similar to GB:L33464 SP:P47734 PID:496118 percent
                identity: 74.23; identified by sequence similarity;
                putative"
                /codon_start=1
                /transl_table=11
                /product="alcohol dehydrogenase, zinc-containing"
                /protein_id="AAF12293.1"
                /db_xref="GI:6460587"
                /translation="MKAIVWQGVNKVGVETVPDPTLLLPTDAIVKITSTAICGSDLHL
                LDGYIPSMETGDILGHEFMGEVVEVGSEVRKLKVGDRVVVPFNIACGVCDPCKRGFFS
                ACDNSNPNHRMAEALYGATSGGALFGYSHMYGGYAGGQAQYVRVPFADVGPFKIESGL
                KDEQVLFLTDIFPTGYQAAEQCNIVPGRDVVAVFGAGPVGQFAARSAQMLGAAHVIVI
                DRVPERLAMAEAAGCQTINYEKEDVLLSLREATGGRGPDHVIDAVGMEAHGHGPGSTM
                DKVQQNLKLTFDRITALRWAILSCAKGGTVSMPGVYGGLVDKMPIGAAFAKGLIFRMG
                QTHTHRHIPSLLGRIEAGQIDPSFVITHRASLDQAPELYKTFRDKQDCCIKVVLNPWA
                "
     gene       complement(5894. .6700)
                /gene="DRA0006"
     CDS        complement(5894. .6700)
                /gene="DRA0006"
                /note="similar to PID:1653488 percent identity: 48.46;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAF12292.1"
                /db_xref="GI:6460586"
                /translation="MDRRSSPQLRTMTNWNREQPAPHAGATGPNSSPGATMQPLERSL
                TGAVGAALAAAGLRSRSGLEKVVLTGLGAGLIAVAARGQNPLATALKIEQNDEGETVV
                RDAVTIGKPAEQLYAVWRDLPGLPLLMTHLRSVEVLDDKRSRWTVEAPAPLGTVSWEA
                ELTADEPGKRIAWRSLPGARIENSGEVLFRPAPGARGTEVVVRLTYRPPGGSAGAVIA
                RMFNQEPSQQLRDDLMRFKREQELGHAPTAQGQSSGRAALTPPNKRGDLV"
     gene       6839. .7744
                /gene="DRA0007"
     CDS        6839. .7744
                /gene="DRA0007"
                /note="similar to GB:AL009126 percent identity: 62.98;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter, ATP-binding protein, putative"
                /protein_id="AAF12291.1"
                /db_xref="GI:6460585"
                /translation="MITTFEQVSKTYGHVTALSDFNLTLRTGELTALLGPNGAGKSTA
                IGLLLGLSAPSAGQVRVLGADPRRNDVRARIGAMPQESALPAGLTVREAVTLFASFYP
                APLGVDEALALADLGPVAGRRAAQLSGGQKRRLAFALAVVGDPELLLIDEPTTGMDAQ
                SRAAFWEAVTGLRARGRTILLTTHYLEEAERTADRVVVMNGGRILADDTPQGLRSGVG
                GARVSFVSDLVQAELERLPGVSAVQVDAAGRADLRTSVPEALLAALIGSGTTFSDLEV
                RRATLEEAYLQLTGPQDMTAVTRSA"
     gene       7861. .8652
                /gene="DRA0008"
     CDS        7861. .8652
                /gene="DRA0008"
                /note="similar to GB:AL009126 percent identity: 52.46;
                identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAF12290.1"
                /db_xref="GI:6460584"
                /translation="MTTLPNVPARRTSALPVLAQLTLGELRRLLRSPMFTVGAIGFPV
                MFFALFGLPAVQEYGATDPHVGPVILTQFAAYSLLSLALFSFGAAVATERSGGWLRLL
                RSSPLPVPLYFVSKTLAALAFGAVSLALLYAFAHFAGGVTLPLGLALLLAGKLLLGMI
                PLVALGLCIGFLASPQAAQILANILSVVMSFASGLFVPLDQLPGFVQQIAPLLPAYHV
                SQIATNTVSGQTASEPAHWLALAAFTLVFGTLAVWGLKRDESREG"
     gene       8809. .9819
                /gene="DRA0009"
     CDS        8809. .9819
```

```
                     /gene="DRA0009"
                     /note="similar to GB:AL009126 percent identity: 55.80;
                     identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAF12289.1"
                     /db_xref="GI:6460583"
                     /translation="MLFWGTLVAFVAVYVFTFGKRDPDDRWAYVGWLSALLMWALGDW
                     LYGPSALSFLVYGGSLIGWQRRQWVALGGAAINALLMVGRLLTLPERSQGDFIFLLFI
                     LASAYANHATYRGLRARRRLAQVQLEKEKLAADAERERIARDLHDLLGHTLSVIVLKS
                     ELAGKLAEKHPQRAAAEIREVERISREALSEVRAAVQGYRGSGLAAELARAKVALDAA
                     GIRLVITDTLPELPADLESTLAMVLREAVTNVVRHSGASELRLTLRGQANGFQLTLQD
                     DGRGGDAPEGTGLNSMRERLRAMGGTFERDGHTGTRLQAFVPLSAGTAGLKAPELLEK
                     TS"
     gene            9816. .10451
                     /gene="DRA0010"
     CDS             9816. .10451

  Query Match           37.6%;  Score 336.2;  DB 1;  Length 213732;
  Best Local Similarity  66.8%;  Pred. No. 4e-67;
  Matches  504;  Conservative    0;  Mismatches  233;  Indels   18;  Gaps    1;

Qy     120 ATGATGTCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTC 179
           | || || | |||||||||||| ||    || || ||  | |||||||| ||||| || |
Db  161202 AAGACGTGAGCATCTACTACGGCGACAAACAGGCCGTCAAAAACGTGAATCTCGATGTGC 161261

Qy     180 CTGCACGCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCC 239
             |  || | |||| ||| |||||||| || || ||||| |||||| | ||  | || |
Db  161262 GCCCTGGCACCGTCAACGCCTTCATCGGGCCGTCGGGCTGCGGCAAGACGACCTTCCTGC 161321

Qy     240 GCTCCATCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTC 299
           |  ||||||| || |||||||||  | ||||| || ||   ||| |  |||   ||||| |
Db  161322 GGGCCATCAACCGCATGCACGACCTGACCCCCGGCGCCCGCGTGACCGGCAGCATCCTGC 161381

Qy     300 TCGACGGCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCG 359
           ||||||| || |||||||| || | |   | ||||| || ||  | ||   |  | | |
Db  161382 TCGACGGGCAGGACATCTACAGCCCGAGCGTGGACCCCGTCGCCATGCGCCGCCGCGTGG 161441

Qy     360 GCATGGTCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTG 419
           | ||||| || || ||| | ||||| ||||| ||||||  | |    ||||||||||| |
Db  161442 GTATGGTGTTTCAAAAGCCCAACCCCTTCCCGACCATGAGCGTGTTCGACAACGTGGTGG 161501

Qy     420 CAGGTCTGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGT 479
           | || ||||| ||  |||||| ||| |||    |  | || | | | || || |||   |
Db  161502 CGGGCCTGAAACTGACCGGCGTGAAAAACCGTGACCACCTGATGCAGGTCGCCGAGCGCT 161561

Qy     480 CTCTTCGTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCG 539
           | || || || ||    |||||||| ||||| || ||||| ||| ||| || | |  ||
Db  161562 CCCTGCGCGGGGCCGCGCTGTGGGACGAGGTCAAAGATCGCCTGAACACCCCCGCCACCG 161621

Qy     540 GCCTCTCCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAG 599
           ||||||| || || |||||||||| |||||||||| |||||| | || || ||||| |
Db  161622 GCCTCTCGGGCGGGCAGCAGCAGCGCCTGTGCATCGCCCGCGCGCTTGCCGTGGAGCCTG 161681

Qy     600 AAATCCTCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGG 659
           |||| || || |||||||| ||   ||| ||||| |||||   |   |||       | |
Db  161682 AAATTCTGCTGATGGACGAACCGACCTCGGCGCTCGACCCTGCCAGCACCGCCAAGATCG 161741

Qy     660 AGGACCTTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGC 719
           |||||||  ||  ||| ||||| |||    ||||||| ||||||||||||||||||||||
Db  161742 AGGACCTGCTCGGCGACCTGAAAAAAGTCACCACCATCATCATCGTGACCCACAACATGC 161801

Qy     720 AGCAGGCTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGC 779
           | ||||| || || ||||| ||    |||  |||| |   ||
Db  161802 ACCAGGCGGCGCGCGTGTCGGACACCACCAGCTTCTTCCTGGTG---------------- 161845

Qy     780 CAGGTCGTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGG 839
             ||    ||||| ||| |||| || ||| |  || | |||   ||||      |  ||
Db  161846 --GGCGACCTGGTCGAACACGGCCCCACCGATCAGGTGTTCACCAACCCGCGCGACGAGC 161903

Qy     840 AAACCGAAGATTACATCTCCGGCCGCTTCGGATAA 874
             ||||| |  ||| | || || || |||||||

Db 161904 GCACCGAGGCCTACGTGTCGGGGCGGTTCGGATGA 161938


RESULT 14
PSEPA1G
LOCUS       PSEPA1G                 3234 bp    DNA     linear   BCT 07-FEB-1999
DEFINITION  Pseudomonas aeruginosa gene.
ACCESSION   D28587
VERSION     D28587.1  GI:465399
KEYWORDS    .
SOURCE      Pseudomonas aeruginosa (strain:PAO1) DNA.
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (bases 1 to 3234)
  AUTHORS   Kato,J., Sakai,Y., Nikata,T. and Ohtake,H.
  TITLE     Cloning and characterization of a Pseudomonas aeruginosa gene
            involved in the negative regulation of phosphate taxis
  JOURNAL   J. Bacteriol. 176 (18), 5874-5877 (1994)
  MEDLINE   94364974
REFERENCE   2  (bases 1 to 3234)
  AUTHORS   Ohtake,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-FEB-1994) Hisao Ohtake, Hiroshima University,
            Fermentation Technology; Kagamiyama 1-4-1, Higashi-Hiroshima,
            Hiroshima 739, Japan (E-mail:hohtake@ue.ipc.hiroshima-u.ac.jp,
            Tel:0824-24-7756, Fax:0824-22-3758)
COMMENT     Submitted (23-Feb-1994) to DDBJ by:
            Hisao Ohtake
            Fermentation Technology
            Hiroshima University
            1-4-1 Kagamiyama, Higashi-Hiroshima
            Hiroshima 724
            Japan
            Phone: 0824-24-7756
            Fax:   0824-22-3758.
FEATURES             Location/Qualifiers
     source          1. .3234
                     /organism="Pseudomonas aeruginosa"
                     /strain="PAO1"
                     /db_xref="taxon:287"
BASE COUNT      608 a   1047 c   1035 g    544 t
ORIGIN

  Query Match           32.3%;  Score 289.2;  DB 1;  Length 3234;
  Best Local Similarity  64.0%;  Pred. No. 1.9e-56;
  Matches  479;  Conservative    0;  Mismatches  248;  Indels   21;  Gaps    2;

Qy     126 TCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCCTGCAC 185
           | ||| | | |||||| |     || ||  |     ||||     |  | ||||    |
Db    1512 TGAACCTGTTCTACGGCGCCAAGCAGGCCCTTTTCGACGTCCGGATGAACATTCCGAAGC 1571

Qy     186 GCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCGCTCCA 245
             || ||||| |||||||| || |||||||| |||||||| ||  | || ||||  |
Db    1572 AGCGCGTGACCGCCTTCATCGGCCCGTCCGGCTGCGGCAAGTCGACCTTGCTGCGCTGCT 1631

Qy     246 TCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCTCGACG 305
           |||| || ||| ||||  |       ||      |||| | ||||||||  |  |||||  |
Db    1632 TCAACCGGATGAACGACCTGGTGGATGGCTGCCGCGTCGAAGGCGAGATTCGCCTCGATG 1691

Qy     306 GCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGGCATGG 365
           || | |||||||  || |       |||||   |  |   | ||      | |||||||||
Db    1692 GCCACAACATCTTCGCCAAGGGCGTCGACGTCGCCGAGCTGCGCCGTCGCGTCGGCATGG 1751

Qy     366 TCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGCAGGTC 425
           | ||||||||| | |||||  |||| |  ||||  | || ||||||||    || |
Db    1752 TGTTCCAGAAGCCCAACCCGTTCC---CCAAGAGCATCTACGAGAACGTGGTCTACGGCC 1808

Qy     426 TGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTCTCTTC 485
           ||    |   |||   || || |||       ||  | || |  |  |||  | | ||
Db    1809 TGCGTATCCAGGGCATCAACAAGAAGCGCGTACTTGACGAGGCCGTCGAGTGGGCGCTGA 1868

Qy     486 GTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGGCCTCT 545
```

```
Db   1869 AGGGGGCTGCGCTGTGGGAGGAGGTGAAGGATCGCCTGCACGAGTCCGCCCTCGGCCTGT 1928

Qy    546 CCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGAAATCC 605

Db   1929 CCGGCGGCCAGCAGCAGCGCCTGGTGATCGCCCGGACCATCGCGGTGGAACCGGAAGTGC 1988

Qy    606 TCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGAGGACC 665

Db   1989 TGCTGCTCGACGAGCCGTGCTCGGCGCTGGACCCGATCTCCACGCTGAAGATCGAAGAGC 2048

Qy    666 TTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCAGCAGG 725

Db   2049 TGATCTACGAGTTGAAGTCGAAGTTCACCATCGTCATCGTTACCCACAACATGCAGCAGG 2108

Qy    726 CTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCCAGGTC 785

Db   2109 CCGCGCGGGTTTCCGACTATACGGCGTTCATGTACATG----------------GGCA 2150

Qy    786 GTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGAAACCG 845

Db   2151 AGCTGATCGAGTTCGGCGACACCGATACGCTGTTCACCAACCCGGCGAAGAAGCAGACCG 2210

Qy    846 AAGATTACATCTCCGGCCGCTTCGGATA 873

Db   2211 AAGACTACATCACCGGCCGCTACGGCTA 2238


RESULT 15
PSEPSTD
LOCUS       PSEPSTD                 6125 bp    DNA     linear   BCT 10-FEB-1999
DEFINITION  Pseudomonas aeruginosa pstC, pstA, pstB and phoU genes (phosphate
            specific transport (pts) operon), complete cds.
ACCESSION   D45195
VERSION     D45195.1  GI:1369910
KEYWORDS    a component of the phosphate specific transport complex; a
            transmembrane component of the phosphate specific transport
            complex; negative regulator of pho regulon; pstC; pstB; pstA; phoU.
SOURCE      Pseudomonas aeruginosa (strain:PAO1) DNA.
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (bases 1 to 6125)
  AUTHORS   Nikata,T., Sakai,Y., Shibat,K., Kato,J., Kuroda,A. and Ohtake,H.
  TITLE     Molecular analysis of the phosphate-specific transport (pst) operon
            of Pseudomonas aeruginosa
  JOURNAL   Mol. Gen. Genet. 250 (6), 692-698 (1996)
  MEDLINE   96204509
REFERENCE   2  (bases 1 to 6125)
  AUTHORS   Ohtake,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-1995) Hisao Ohtake, Hiroshima University,
            Fermentation Technology; Kagamiyama 1-4-1, Higashi-Hiroshima,
            Hiroshima 739, Japan (E-mail:hohtake@ue.ipc.hiroshima-u.ac.jp,
            Tel:0824-24-7756, Fax:0824-22-3758)
COMMENT     On Jun 8, 1996 this sequence version replaced gi:633991.
FEATURES             Location/Qualifiers
     source          1. .6125
                     /organism="Pseudomonas aeruginosa"
                     /strain="PAO1"
                     /db_xref="taxon:287"
     protein_bind    259. .276
                     /standard_name="pho box"
                     /bound_moiety="PhoB"
                     /function="PhoB binding to pho box leads to activation of
                     transcription"
     gene            553. .2586
                     /gene="pstC"
     CDS             553. .2586
                     /gene="pstC"
                     /codon_start=1
                     /transl_table=11
                     /product="a transmembrane component of the phosphate
                     specific transport complex"
                     /protein_id="BAA08135.1"
                     /db_xref="GI:633992"
                     /translation="MQQAQVPLLLSIEEQNQVGMRVGGQGEILFFNAKTGVELKRVAL
                     PLPAGSEIRSIAQDQPGSPTVALGLSNGQVLVFQHTYKVTYPDNKKTITPEIAFPYGE
                     TPIGLDPQGRPLEHVSINAGDDSLLLAGSVDKQLLLLSMTREENMLTGESTLDEERIE
                     LPQIAEPVKAIYLDPRKQWLYVINGRATADVFDLHSRQLNGRYKLLEDPNAEVTASTQ
                     LLGGISLLVGDSKGGIAQWFMARDTDGEPRLSHVRDFNLDGAPISAIAPEQRRKGFIA
                     LDEKGNLGVFHSTAHRTLLVQPVADSSGVITLSPRANRLLLEQGGKIHRFALSNPHPE
                     ISWSALWGKVWYESYDKPAYVWQSTAATTDFEPKLSLSPLTFGTLKAAFYAMILAAPL
                     AIAAAVYTAYFMAPAMRRKVKPVIELMEALPTVILGFFAGLFLAPYVEGHLPGVFSLL
                     LLTPLGILLAGLLWSRLPERIRLSLPAGWEAALLIPVVLAVGAFALWLSPHLETTFFG
                     GDMRLWLSHDLGITYDQRNALIVGLAMGFAVIPNIFSIAEDAIFSVPRSLTYGSLALG
                     ATPWQTLTRVVILTASPGIFSALMIGMGRAVGETMIVLMATGNTPVMEVNIFEGMRTL
                     AANVAVEMPESEVGGSHYRVLFLSALVLLTFTFVMNTLAELIRQRLRKKYASL"
     gene            2666. .4207
                     /gene="pstA"
     CDS             2666. .4207
                     /gene="pstA"
                     /codon_start=1
                     /transl_table=11
                     /product="a transmembrane component of the phosphate
                     specific transport complex"
                     /protein_id="BAA08136.1"
                     /db_xref="GI:1369911"
                     /translation="MNAGAVSIAVIMTIGLLAVIAVRGLGHFWPANVIEATYQIPGEA
                     AKTVIGEVVQAEEVPRTRLRSAGMPVPEAGPEFMTRELLKLGNRELYGSDFSWVVGEW
                     LQDIKRPEALMVLERREWGNFYGYLLSVKENGQVVAEGDAAVGELQKRMQRVDDLYAK
                     LYKLEKGDIGGINHGLERLRLKERKLQLDGKLDAAAQADLAAERSELQARYKVLEEQL
                     NGLHQEFNRDSVVMRDASGRESEITLGKLVHAYQPNAMGLGTKMTVYFKKLWEFLSDD
                     PREANTEGGIFPAIFGTVMMTLVMALIVTPFGVIAAVYLREYAKQGPLTRVIRIAVNN
                     LAGVPAIVYGVFGLGFFVYVLGGSIDRLFPEALPAPTFGTPGLFWASLTLAILAVPV
                     VIVATEEGLARIPRAVREGSLALGATKAETLWKIVLPMASPAMMTGLILAVARAAGEV
                     APLMLVGVVKLAPSLPVDGNYPYVHLDQKIMHLGFHIYDVGFQSPNVEAARRWYTLRR
                     CCWSW"
     gene            4297. .5130
                     /gene="pstB"
     CDS             4297. .5130
                     /gene="pstB"
                     /codon_start=1
                     /transl_table=11
                     /product="a component of the phosphate specific transport
                     complex"
                     /protein_id="BAA08137.1"
                     /db_xref="GI:633995"
                     /translation="MQNETASHGINFDALGRDRQSLDLASESVELEVPGLNLFYGAKQ
                     ALFDVRMNIPKQRVTAFIGPSGCGKSTLLRCFNRMNDLVDGCRVEGEIRLDGHNIFAK
                     GVDVAELRRRVGMVFQKPNPFPKSIYENVVYGLRIQGINKKRVLDEAVEWALKGAALW
                     EEVKDRLHESALGLSGGQQQRLVIARTIAVEPEVLLLDEPCSALDPISTLKIEELIYE
                     LKSKFTIVIVTHNMQQAARVSDYTAFMYMGKLIEFGDTDTLFTNPAKKQTEDYITGRY
                     G"
     gene            5226. .5954
                     /gene="phoU"
     CDS             5226. .5954
                     /gene="phoU"
                     /codon_start=1
                     /transl_table=11
                     /product="a negative regulator of pho regulon"
                     /protein_id="BAA08138.1"
                     /db_xref="GI:633996"
                     /translation="MINKDSLTHHISQQFNAELEDVRSHLLAMGGLVEKQVNDAVNAL
                     IDADSGLAQQVREIDDQINQMERNIDEECVRILARRQPAASDLRLIISISKSVIDLER
                     IGDEASKVARRAIQLCEEGESPRGYVEVRHIGSQVQKMVQEALDAFARFDADLALSVA
                     QYDKTVDREYKTALRELVTYMMEDPRAISRVLNIIWALRSLERIGDHARNIAELVIYL
                     VRGTDVRHIGLTRMKEEVENNRGE"
BASE COUNT     1091 a   2042 c   1949 g   1043 t
ORIGIN

  Query Match             32.3%;  Score 289.2;  DB 1;  Length 6125;
  Best Local Similarity   64.0%;  Pred. No. 2.1e-56;
  Matches  479;  Conservative    0;  Mismatches  248;  Indels   21;  Gaps    2;

Qy    126 TCAACATCTACTACGGTGATTTCCACGCAGTGCAGAACGTGAACCTCGAGGTTCCTGCAC 185
```

```
Db   4403 TGAACCTGTTCTACGGCGCCAAGCAGGCCCTTTTCGACGTCCGGATGAACATTCCGAAGC 4462

Qy    186 GCTCTGTCACCGCATTCATCGGACCATCCGGCTGTGGCAAGTCCACAGTTCTCCGCTCCA 245

Db   4463 AGCGCGTGACCGCCTTCATCGGCCCGTCCGGCTGCGGCAAGTCGACCTTGCTGCGCTGCT 4522

Qy    246 TCAATCGTATGCACGAGGTCACCCCAGGTGCATACGTCAAGGGCGAGATCCTTCTCGACG 305

Db   4523 TCAACCGGATGAACGACCTGGTGGATGGCTGCCGCGTCGAAGGCGAGATTCGCCTCGATG 4582

Qy    306 GCGAGAACATCTACGGCTCCAAGATCGACCCAGTTGCAGTCCGTAACACCATCGGCATGG 365

Db   4583 GCCACAACATCTTCGCCAAGGGCGTCGACGTCGCCGAGCTGCGCCGTCGCGTCGGCATGG 4642

Qy    366 TCTTCCAGAAGGCTAACCCATTCCCAACCATGTCCATCGAGGACAACGTGGTTGCAGGTC 425

Db   4643 TGTTCCAGAAGCCCAACCCGTTCC---CCAAGAGCATCTACGAGAACGTGGTCTACGGCC 4699

Qy    426 TGAAGCTTTCCGGCGAGAAGAACAAGAAGAAGCTCAAGGAAGTTGCTGAGAAGTCTCTTC 485

Db   4700 TGCGTATCCAGGGCATCAACAAGAAGCGCGTACTTGACGAGGCCGTCGAGTGGGCGCTGA 4759

Qy    486 GTGGCGCAAACCTGTGGGAAGAGGTTAAGGATCGTCTGGACAAGCCAGGCGGCGGCCTCT 545

Db   4760 AGGGGGCTGCGCTGTGGGAGGAGGTGAAGGATCGCCTGCACGAGTCCGCCCTCGGCCTGT 4819

Qy    546 CCGGTGGTCAGCAGCAGCGTCTGTGCATCGCTCGCGCGATCGCGGTTGAGCCAGAAATCC 605

Db   4820 CCGGCGGCCAGCAGCAGCGCCTGGTGATCGCCCGGACCATCGCGGTGGAACCGGAAGTGC 4879

Qy    606 TCCTCATGGACGAGCCTTGCTCCGCGCTTGACCCAATCTCAACCCTGGCTGTGGAGGACC 665

Db   4880 TGCTGCTCGACGAGCCGTGCTCGGCGCTGGACCCGATCTCCACGCTGAAGATCGAAGAGC 4939

Qy    666 TTATCCACGAGCTGAAGGAAGAGTTCACCATCGTCATCGTGACCCACAACATGCAGCAGG 725

Db   4940 TGATCTACGAGTTGAAGTCGAAGTTCACCATCGTCATCGTTACCCACAACATGCAGCAGG 4999

Qy    726 CTGCACGTGTGTCCGATCAGACCGCGTTCTACTCCCTGGAGGCGACCGGTAGGCCAGGTC 785

Db   5000 CCGCGCGGGTTTCCGACTATACGGCGTTCATGTACATG----------------GGCA 5041

Qy    786 GTCTGGTTGAAATCGGACCTACCAAGAAGATCTTCGAAAACCCAGATCAGAAGGAAACCG 845

Db   5042 AGCTGATCGAGTTCGGCGACACCGATACGCTGTTCACCAACCCGGCGAAGAAGCAGACCG 5101

Qy    846 AAGATTACATCTCCGGCCGCTTCGGATA 873

Db   5102 AAGACTACATCACCGGCCGCTACGGCTA 5129


Search completed: April 24, 2003, 16:49:41
Job time : 3499 secs
```